

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:13:34 ; Search time 163.5 Seconds  
(without alignments)  
4056.844 Million cell updates/sec

Title: US-10-021-571-4  
Perfect score: 8884  
Sequence: 1 MSVLISQSVINYVEENIPA.....LHAASSESTGFGBERESIL 1715

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8884	100.0	1715	4	AAM39025 Human pol
2	8884	100.0	1715	4	AAM38993 Human pol
3	8884	100.0	1715	5	AAE25144 Human ARM
4	8884	100.0	1715	5	AAU96840 Human kid
5	8884	100.0	1715	6	AAE32128 Human cyt
6	8855	99.7	1753	6	ADA09888 Human rec
7	8830	99.4	1771	7	ADC06847 Human Kid
8	8362.5	94.1	1762	5	AAU96841 Rat Kidin
9	8362.5	94.1	1762	7	ADC06848 Rat Kidin
10	8301	93.4	1763	5	AAU80244 Rat Kidin
11	8266	93.0	1715	5	AAE25143 Rat ARMS
12	8266	93.0	1715	7	ADC06849 Rat Kidin
13	5931	66.8	1142	7	ADC06850 Kidins 22
14	5899.5	66.4	1184	7	ADC06846 Human Kid
15	3050.5	34.3	705	4	AAU75604 Human col
16	2851.5	32.1	551	4	AAE01035 Human dea
17	2791	31.4	543	4	AAAB95191 Human pro
18	2638	29.7	513	4	AAE03645 Human ext
19	2503	28.2	624	4	ABG08697 Novel hum
20	2379.5	26.8	1498	4	ABB64857 Drosophil
21	2379.5	26.8	1498	5	AAE25146 Fruit fly
22	1832.5	20.6	1398	5	AAE25145 Worn ARMS
23	667	7.5	131	5	ADK36597 Novel hum
24	666	7.5	129	4	AAAB94785 Human pro
25	620	7.0	120	4	AAAM40811 Human pol

ALIGNMENTS

RESULT 1  
AAM39025  
ID AAM39025 standard; protein; 1715 AA.

XX AC AAM39025;  
XX XX  
DT 22-OCT-2001 (first entry)  
DE Human polypeptide SEQ ID NO 2170.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX FD 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.  
XX 23-DEC-1999; 99US-00471275.  
XX 21-JAN-2000; 2000US-00488725.  
XX 25-APR-2000; 2000US-00552317.  
XX 20-JUN-2000; 2000US-00598042.  
XX 19-JUL-2000; 2000US-00620312.  
XX 03-AUG-2000; 2000US-00653450.  
XX 14-SEP-2000; 2000US-00662191.  
XX 19-OCT-2000; 2000US-00693036.  
XX 29-NOV-2000; 2000US-00727344.  
(HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI58181.  
PT Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX AAB94785 Human pro  
PS Example 4; SEQ ID NO 2170; 10078pp; English.

Add47763 Human pro  
Aab56277 Human sec  
Aad90358 Full leng  
Abg31785 Human ank  
Aac39122 Novel hum  
Aae33667 Human str  
Adb64375 Human pro  
Aam83911 Human dia  
Aad09291 Human pro  
Aam83918 Human dia  
Aam83919 Human dia  
Aam83916 Human dia  
Aam83915 Human dia  
Aam83914 Human dia  
Aam83912 Human dia  
Aam83913 Human dia  
Aam83917 Human dia  
Add27862 Human bra  
Aam79160 Human pro  
Abg00972 Novel hum

XX	The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA38642-AA42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemia and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification									
XX	SQ Sequence 1715 AA;									
	Query Match	100.0%;	Score 8884;	DB 4;	Length 1715;					
	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 1715;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
QY	1	MSVLISQSVINYYEENIPALKALLECKDQVDERNECGQTPLMIAAEQGNLEIVKELIKN	60							
DB	1	MSVLISQSVINYYEENIPALKALLECKDQVDERNECGQTPLMIAAEQGNLEIVKELIKN	60							
QY	61	GANCNLEDLNTALISASKEGHVHIVBELLKCGVNLHRDMGWTALMWACYKGRTDVV	120							
DB	61	GANCNLEDLNTALISASKEGHVHIVBELLKCGVNLHRDMGWTALMWACYKGRTDVV	120							
QY	121	ELLISHGANPSVTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNSCDKYGTTPLVAAAR	180							
DB	121	ELLISHGANPSVTGLQYSVYPIIWAAGRGHADIHVLHLLQNGAKVNSCDKYGTTPLVAAAR	180							
QY	181	KGHLECVKHLAMGADVQDQGANSMATLIIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA	240							
DB	181	KGHLECVKHLAMGADVQDQGANSMATLIIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA	240							
QY	241	LMIASKEGHTETIIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	300							
DB	241	LMIASKEGHTETIIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	300							
QY	301	QDNKTALYWAVEKGNATWDRDILQCPDTEICTKDGTEPLIKATKRNTEVWELLDDKGA	360							
DB	301	QDNKTALYWAVEKGNATWDRDILQCPDTEICTKDGTEPLIKATKRNTEVWELLDDKGA	360							
QY	361	KVSAVDKKGDTPLHIAIRGRSRKLAELLRNPKDGRLLYRPNKAGETPPYINIDCSHOKSIL	420							
DB	361	KVSAVDKKGDTPLHIAIRGRSRKLAELLRNPKDGRLLYRPNKAGETPPYINIDCSHOKSIL	420							
QY	421	TOIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKKL	480							
DB	421	TOIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQMGSGKSFLLKKL	480							
QY	481	EDEMTFAGQIQEPLFOFSLIIVFLTLILCCGGLLPAFTVHNPLGIVLSFLALLIYF	540							
DB	481	EDEMTFAGQIQEPLFOFSLIIVFLTLILCCGGLLPAFTVHNPLGIVLSFLALLIYF	540							
QY	541	FIVYFGREGESWNWVLSRLARHIGLYELLKLMFVNPPPELPEOTTKALPVRFPLF	600							
DB	541	FIVYFGREGESWNWVLSRLARHIGLYELLKLMFVNPPPELPEOTTKALPVRFPLF	600							
QY	601	TDYNRLSSVGGETSIAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKKTCCPLS	660							
DB	601	TDYNRLSSVGGETSIAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKKTCCPLS	660							
QY	661	FVIFLFIIGCIISGITLLAIFRVDPKHLTWNVLISIASVVGLAFLVLCNRTWQVLDLSLL	720							
DB	661	FVIFLFIIGCIISGITLLAIFRVDPKHLTWNVLISIASVVGLAFLVLCNRTWQVLDLSLL	720							
QY	721	NSORKELHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIIDTSFTQNTQRLVLIIDGLDAC	780							

RESULT 2

AA388993

ID AA388993 standard; protein; 1715 AA.

XX

AC AA388993;

XX

DT 22-OCT-2001 (first entry)  
XX Human polypeptide SEQ ID NO 2138.  
XX  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.  
XX  
XX Homo sapiens.  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US034263.  
XX  
XX 23-DEC-1999; 99US-00471275.  
XX 21-JAN-2000; 2000US-00488725.  
XX 25-APR-2000; 2000US-00552317.  
XX 20-JUN-2000; 2000US-00598042.  
XX 19-JUL-2000; 2000US-00620312.  
XX 03-AUG-2000; 2000US-00653450.  
XX 14-SEP-2000; 2000US-00662191.  
XX 19-OCT-2000; 2000US-00693036.  
XX 29-NOV-2000; 2000US-00727344.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
XX N-PSDB; AA158149.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.  
XX  
XX Example 4; SEQ ID NO 2138; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA139642-AA42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemia and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification  
XX  
XX Sequence 1715 AA;  
XX  
XX Query Match 100.0%; Score 8884; DB 4; Length 1715;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 MSVLSQSVINYVEENIPALKALLECKVDNERNECCQTFLMTAAEQNLEIVKELIKN 60  
XX  
XX 1 MSVLSQSVINYVEENIPALKALLECKVDNERNECCQTFLMTAAEQNLEIVKELIKN 60  
XX  
XX 61 GANCNLEDLDNWTALISAKSGHVHIVBELKCGVNLNLEHROMGWWTALMWACYKGRDGV 120  
XX  
XX 61 GANCNLEDLDNWTALISAKSGHVHIVBELKCGVNLNLEHROMGWWTALMWACYKGRDGV 120  
XX  
XX 1201 KEMMNFQDWHLFRSTVLEMRNAESHVVPDRFLSESSSGPAPHGEPARRASHNELPHT 1260

QY 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPVWAAR 180  
DB 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPVWAAR 180  
QY 181 KGHLECVKHLAMGADVQEGANSWTALIIVAVKGYTOSVKEILKRNPNVNLTDKGNFA 240  
DB 181 KGHLECVKHLAMGADVQEGANSWTALIIVAVKGYTOSVKEILKRNPNVNLTDKGNFA 240  
QY 241 LMIAKSGHTEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300  
DB 241 LMIAKSGHTEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300  
QY 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKORNIIEVVELLDKGA 360  
DB 301 QDNKTALYWAVEKGNATMVRDILQCNPDTEICTKDGETPLIKATKORNIIEVVELLDKGA 360  
QY 361 KVSADVKKGDTPLHIAIRGRSRKLAELLLRNPDKGRLLYRPNKAGETPINIDCSHQSIL 420  
DB 361 KVSADVKKGDTPLHIAIRGRSRKLAELLLRNPDKGRLLYRPNKAGETPINIDCSHQSIL 420  
QY 421 TOIFGARHLSPTETDGMGLYDLYSSALADILSPTMOPPICVGIYAQWGSKSLKKL 480  
DB 421 TOIFGARHLSPTETDGMGLYDLYSSALADILSPTMOPPICVGIYAQWGSKSLKKL 480  
QY 481 EDEMKTFAQQIEPLFQFQSWLIVFLTLILCGGLGLLFAFTVHPNLGIAVSLSFLALLYIF 540  
DB 481 EDEMKTFAQQIEPLFQFQSWLIVFLTLILCGGLGLLFAFTVHPNLGIAVSLSFLALLYIF 540  
QY 541 FIVYFGRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPPEQTTKALPVRLF 600  
DB 541 FIVYFGRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPPEQTTKALPVRLF 600  
QY 601 TDYNRLSVGGTSLAEMIATLSDACRECFPLATRLPRVFKETDTCQKKWKKTCCCLPS 660  
DB 601 TDYNRLSVGGTSLAEMIATLSDACRECFPLATRLPRVFKETDTCQKKWKKTCCCLPS 660  
QY 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTNAVLISIASVVGGLAVLNCRTWQVLDLIL 720  
DB 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTNAVLISIASVVGGLAVLNCRTWQVLDLIL 720  
QY 721 NSQRKRLHNAASKLHKLKSEGFMKLVCEVELMARMAKTIDSFQNTQRLVVIIDGLDAC 780  
DB 721 NSQRKRLHNAASKLHKLKSEGFMKLVCEVELMARMAKTIDSFQNTQRLVVIIDGLDAC 780  
QY 781 EODKVLQMLDTRVLFSGKPPFIAIFASDPHIIKAINQNLNSVLDRDSNINGHDYMRNVH 840  
DB 781 EODKVLQMLDTRVLFSGKPPFIAIFASDPHIIKAINQNLNSVLDRDSNINGHDYMRNVH 840  
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDVTGIEQADRRVSONSLGEMTKLGSKTALN 900  
DB 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDVTGIEQADRRVSONSLGEMTKLGSKTALN 900  
QY 901 RRDYRRRQMQRTITRQMSFDLTCLLVTEWDFSDISQPMRRLNIVSVTGRLLRANQIS 960  
DB 901 RRDYRRRQMQRTITRQMSFDLTCLLVTEWDFSDISQPMRRLNIVSVTGRLLRANQIS 960  
QY 961 FNDRLASWINLTQEWPYRTSWLILYLEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020  
DB 961 FNDRLASWINLTQEWPYRTSWLILYLEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020  
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIIVADVRAAREQISIGGLAY 1080  
DB 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIIVADVRAAREQISIGGLAY 1080  
QY 1081 PPLPLHGPPRAPSGYQPPSVCSSTSFNGFPAGVSPQPHSSYSGMTGPQHPFYNRG 1140  
DB 1081 PPLPLHGPPRAPSGYQPPSVCSSTSFNGFPAGVSPQPHSSYSGMTGPQHPFYNRG 1140  
QY 1141 SGPRAGPVLLNSLNDVAVCEKLEKQIEGLDOSMLPOYCTTIKKANINGRVLAOCNIDELK 1200  
DB 1141 SGPRAGPVLLNSLNDVAVCEKLEKQIEGLDOSMLPOYCTTIKKANINGRVLAOCNIDELK 1200  
QY 1201 KEMMNFQDWHLFRSTVLEMRNAESHVVPDRFLSESSSGPAPHGEPARRASHNELPHT 1260

```

Db 1201 KENMMFGDWHLFRSTVLENRNASHVVPDPFLSESSGPAHPGEPARRASHNELPHT 1260
QY 1261 ELSSTQPTTLNLFNFEELNTLGLDEGAPRHSNLSQWQTRTPSLSLNSQSSIEISKLT 1320
Db 1261 ELSSTQPTTLNLFNFEELNTLGLDEGAPRHSNLSQWQTRTPSLSLNSQSSIEISKLT 1320
QY 1321 DKVQAEYRDYRYIIAQMQLGEGPGSTTISGRSSPHSTTYMGQSSGGSIHNSLQEK 1380
Db 1321 DKVQAEYRDYRYIIAQMQLGEGPGSTTISGRSSPHSTTYMGQSSGGSIHNSLQEK 1380
QY 1381 KDEPKPDGKGFMLKRGVIVDIYSSGVSTNDASPLDPITEDEKSDQSGSKLLPKKKS 1440
Db 1381 KDEPKPDGKGFMLKRGVIVDIYSSGVSTNDASPLDPITEDEKSDQSGSKLLPKKKS 1440
QY 1441 SERSSLPQTDLKLKGLSLRYQKLPDSESGTEESDNTPLKDDKDKAEGKVERVPKSP 1500
Db 1441 SERSSLPQTDLKLKGLSLRYQKLPDSESGTEESDNTPLKDDKDKAEGKVERVPKSP 1500
QY 1501 EHSABIRTPIKAKAYLSKLLDKQSSDSCVRSSESSPHSLHNEVADDSQLEKANLIE 1560
Db 1501 EHSABIRTPIKAKAYLSKLLDKQSSDSCVRSSESSPHSLHNEVADDSQLEKANLIE 1560
QY 1561 LEDDSHSGKRGIPHSLSGLQDPIIARMSICSEDKSPSECSLIASSPEENWPACQKAYNL 1620
Db 1561 LEDDSHSGKRGIPHSLSGLQDPIIARMSICSEDKSPSECSLIASSPEENWPACQKAYNL 1620
QY 1621 NRTPTSTVLNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQENLKSMTNKR 1680
Db 1621 NRTPTSTVLNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQENLKSMTNKR 1680
QY 1681 SQSSSYRLSKDPPELHAAASSESTGFERESIL 1715
Db 1681 SQSSSYRLSKDPPELHAAASSESTGFERESIL 1715

RESULT 3
AAE25144
ID AAE25144 standard; protein; 1715 AA.
AC AAE25144;
XX
XX
XX 30-OCT-2002 (first entry)
XX Human ARMS protein.
DE
XX
XX Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
XX ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
XX diagnostic; imaging; human.
XX
XX Homo sapiens.
OS
XX
XX Key
XX Location/Qualifiers
XX 17..390
XX /note= "Ankyrin repeat"
XX 354..493
XX /note= "N-terminal region"
XX 496..518
XX /note= "Transmembrane domain"
XX 525..546
XX /note= "Transmembrane domain"
XX 573..638
XX /note= "Loop between TM2 and TM3"
XX 661..680
XX /note= "Transmembrane domain"
XX 688..710
XX /note= "Transmembrane domain"
XX 749..854
XX /note= "Carboxy tail"
XX 940..1060
XX /note= "Carboxy tail"
XX 1081..1093
XX /note= "Polyproline stretch"

```

```

FT Domain 1151..1221
FT Binding-site /note= "SAM domain"
FT 1713..1715
FT /note= "PDZ binding motif"
XX WO200250273-A2.
XX
XX 27-JUN-2002.
XX
XX 19-DEC-2001; 2001WO-US048603.
XX
XX 21-DEC-2000; 2000US-0256909P.
XX
XX (UYN ) UNIV NEW YORK STATE.
XX
XX Chao MV, Kong H;
XX PI
XX WPI; 2002-508800/54.
XX N-PSDB; AAD41036.
XX
XX Isolated ankyrin repeat-rich membrane spanning (ARMS) polypeptide that is
XX a target for phosphorylation by neurotrophin and ephrin receptor tyrosine
XX kinases, useful as a marker for growth cones.
XX
XX Claim 1; Page 103-110; 136pp; English.
XX
XX The invention relates to ankyrin repeat-rich membrane spanning (ARMS)
XX protein which is a down stream target of neurotrophin and ephrin receptor
XX tyrosine kinases. ARMS DNA is useful for visualising the growth cone of
XX neurons. ARMS protein is useful as an indicator of the biological
XX activity of neurotrophins and ephrins, as a marker for neuronal cells
XX which have the ability to undergo continued synaptic changes through
XX adult life or for the presence and distribution of ARMS in such neuronal
XX cells, or as a marker for growth cones. ARMS DNA is useful in diagnostic
XX and imaging methods. The present sequence is human ARMS protein
XX
XX Sequence 1715 AA;
XX
XX Query Match 100.0%; Score 8884; DB 5; Length 1715;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSVLISQSVINVEENIPALKLEKCKDVDERNECGOTPLMIAAEQGNLEIVKELIKN 60
Db 1 MSVLISQSVINVEENIPALKLEKCKDVDERNECGOTPLMIAAEQGNLEIVKELIKN 60
QY 61 GANCNLELDNWTALISASKEGHVHIVEELLKCGVNLHRDMGGWTALMWACYKGRTDVV 120
Db 61 GANCNLELDNWTALISASKEGHVHIVEELLKCGVNLHRDMGGWTALMWACYKGRTDVV 120
QY 121 ELLLSHGANSVTLGQYSVYPIIWAAGRGHADIIVHLLQLONGAKVNSDKYGTTPPLVWAAR 180
Db 121 ELLLSHGANSVTLGQYSVYPIIWAAGRGHADIIVHLLQLONGAKVNSDKYGTTPPLVWAAR 180
QY 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTQSVKELKENPNVNLTKDGNTA 240
Db 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTQSVKELKENPNVNLTKDGNTA 240
QY 241 LMIASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
Db 241 LMIASKEGHEIVQDILLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
QY 301 QNKNTALYWAVKGNATWVRDILQCNPDTEICTKQGETPLIKATKORNIIEVVELLDKGA 360
Db 301 QNKNTALYWAVKGNATWVRDILQCNPDTEICTKQGETPLIKATKORNIIEVVELLDKGA 360
QY 361 KVSADVKKGDTPLHIAIRGRSRKLAELLRNPKDGLLYRPNKAGETPNYIDCSHOKSIL 420
Db 361 KVSADVKKGDTPLHIAIRGRSRKLAELLRNPKDGLLYRPNKAGETPNYIDCSHOKSIL 420
QY 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPIICVGLYAQWNGSGKSFLLKKL 480
Db 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPIICVGLYAQWNGSGKSFLLKKL 480

```



QY 481 EDEMTFAGQOIEPLFQFQSWLIVFLTLCCGGLGLLFAFTVHPNIGIAVSLSFALLYIF 540  
Db 481 EDEMTFAGQOIEPLFQFQSWLIVFLTLCCGGLGLLFAFTVHPNIGIAVSLSFALLYIF 540  
QY 541 FIVIFYGGRREGESNNWAWLSTRARHIGYLELLKLMFVNPPPELPTQTKALPVRPLF 600  
Db 541 FIVIFYGGRREGESNNWAWLSTRARHIGYLELLKLMFVNPPPELPTQTKALPVRPLF 600  
QY 601 TDYNRLSVGGETS LAEMIATLSDACERFGLFATRLFRVFKTEDTQCKKWKTCCLPS 660  
Db 601 TDYNRLSVGGETS LAEMIATLSDACERFGLFATRLFRVFKTEDTQCKKWKTCCLPS 660  
QY 661 FVIFLFIIGCIISGTTLLAI FRVDPKHLTVNAVLISIASVVGGLAFVLCNRTWQVLDL 720  
Db 661 FVIFLFIIGCIISGTTLLAI FRVDPKHLTVNAVLISIASVVGGLAFVLCNRTWQVLDL 720  
QY 721 NSQRKLLNAASKLHLKSEGFMKYLCEVELMARMKTIIDFTQNTQLRVVIIDGLDAC 780  
Db 721 NSQRKLLNAASKLHLKSEGFMKYLCEVELMARMKTIIDFTQNTQLRVVIIDGLDAC 780  
QY 781 EODKVLQMLDTRVFLFSKGPPIAFASDPHIIIIKAINONLNSVLKSDNSINGHDYMRNIVH 840  
Db 781 EODKVLQMLDTRVFLFSKGPPIAFASDPHIIIIKAINONLNSVLKSDNSINGHDYMRNIVH 840  
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGIGQEDADRRVQNSLGEMTKLASKTALN 900  
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGIGQEDADRRVQNSLGEMTKLASKTALN 900  
QY 901 RRDYRRRQMQRTITRQMSFDLTKLLVTEDFWFSDISPOTMRRLLNIVSVTGRLLRANQIS 960  
Db 901 RRDYRRRQMQRTITRQMSFDLTKLLVTEDFWFSDISPOTMRRLLNIVSVTGRLLRANQIS 960  
QY 961 FNMRLASWINLTEOMPVRTSWLILYLEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020  
Db 961 FNMRLASWINLTEOMPVRTSWLILYLEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020  
QY 1021 EIDGDIRNFVFLSSRTFVLVARDVKVFLPCTVNLDPKRLIADVRAAREQISIGGLAY 1080  
Db 1021 EIDGDIRNFVFLSSRTFVLVARDVKVFLPCTVNLDPKRLIADVRAAREQISIGGLAY 1080  
QY 1081 PPLPLEHGPFPAPGYSQPPSCVSTSFNGPFAGVSPQPHSSYSYSGMTQFPHPFVNRG 1140  
Db 1081 PPLPLEHGPFPAPGYSQPPSCVSTSFNGPFAGVSPQPHSSYSYSGMTQFPHPFVNRG 1140  
QY 1141 SGAPAPGVVLLNSLNVDAVCEKIQIEGLDOSMLPOYCTTIKKANINGRVLAQCNIDELK 1200  
Db 1141 SGAPAPGVVLLNSLNVDAVCEKIQIEGLDOSMLPOYCTTIKKANINGRVLAQCNIDELK 1200  
QY 1201 KEMNNFQDGMHLFRSTVLEMEANESHVVVPDPRFLSESSSGPAPHGEPARRASHNELPHT 1260  
Db 1201 KEMNNFQDGMHLFRSTVLEMEANESHVVVPDPRFLSESSSGPAPHGEPARRASHNELPHT 1260  
QY 1261 ELSSQTPYTLNFSFEELNTGLDEGAPRHSNLSWQSQTRRTPSLSSLSNQDSSIEISKLT 1320  
Db 1261 ELSSQTPYTLNFSFEELNTGLDEGAPRHSNLSWQSQTRRTPSLSSLSNQDSSIEISKLT 1320  
QY 1321 DKVQAEYRDAYREYTAQNSQLEGGPGSTTISGRSPHSTYTMGSSSGSSTHSLNLEQKG 1380  
Db 1321 DKVQAEYRDAYREYTAQNSQLEGGPGSTTISGRSPHSTYTMGSSSGSSTHSLNLEQKG 1380  
QY 1381 KDSBPKPDDGRKSLFKRKGDDVIDYSSSGVSTNDASPLDPIITEEKEKSDQSGSKLLPGKKS 1440  
Db 1381 KDSBPKPDDGRKSLFKRKGDDVIDYSSSGVSTNDASPLDPIITEEKEKSDQSGSKLLPGKKS 1440  
QY 1441 SERSSLFOTDLKXGSLRYQKLPSEDESGTEESDNTPLKDKDKRKAEGKVERVPKSP 1500  
Db 1441 SERSSLFOTDLKXGSLRYQKLPSEDESGTEESDNTPLKDKDKRKAEGKVERVPKSP 1500  
QY 1501 EHSAPERTPTFKAEYLSALLDKKSDSDSGVRSESSPNHSLHNEVADDSQLEKANLIE 1560  
Db 1501 EHSAPERTPTFKAEYLSALLDKKSDSDSGVRSESSPNHSLHNEVADDSQLEKANLIE 1560

QY 1561 LEDSHSGKRGIPHSLGLODPIIARMSICSDKSPSECSLIASSPEENWPACQAYNL 1620  
Db 1561 LEDSHSGKRGIPHSLGLODPIIARMSICSDKSPSECSLIASSPEENWPACQAYNL 1620  
QY 1621 NRTSTVTLLNNSAPANRANQNFDEMEGIRTSQVILRPSSSPNPTTTONENLKSMTKTR 1680  
Db 1621 NRTSTVTLLNNSAPANRANQNFDEMEGIRTSQVILRPSSSPNPTTTONENLKSMTKTR 1680  
QY 1681 SQRSSYTRLSDKDPPELHAAASSESTGFGERESIL 1715  
Db 1681 SQRSSYTRLSDKDPPELHAAASSESTGFGERESIL 1715  
RESULT 4  
AAU96840  
ID AAU96840 standard; protein; 1715 AA.  
AC AAU96840;  
XX 30-JUL-2002 (first entry)  
XX Human kidins220 protein.  
XX Kidins220; kinase D interacting substrate of 22kDa; cytostatic;  
KW neuroprotective; gene therapy; protein kinase D; PKD; cancer;  
KW neurodegenerative disease; glioblastoma multiforme; prostate cancer;  
KW human.  
XX Homo sapiens.  
OS WO200220786-A2.  
XX 14-MAR-2002.  
XX 06-SEP-2001; 2001WO-GB003977.  
XX 06-SEP-2000; 2000US-0230449P.  
(IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
Schiaivo G, Iglesias T;  
WPI; 2002-371879/40.  
DR N-PSDB; ABK51221.  
XX Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,  
PT useful for identifying modulators useful in treating cancer and  
PT neurodegenerative diseases.  
XX Claim 33; Fig 11; 228pp; English.  
XX The invention relates to polypeptide comprising a kinase D interacting  
CC substrate of 220 kDa (Kidins220) from rat and human or their fragment,  
CC variant or fusion provided that the protein is not the polypeptide  
CC encoded by GenBank Accession No. AB033076. Also included are a  
CC polynucleotide encoding Kidins220 provided that polynucleotide is not any  
CC one of the clones corresponding to the 61 GenBank Accession Nos given in  
CC the specification; an expression/replicable vector comprising the  
CC polynucleotide; a recombinant host cell containing the polynucleotide or  
CC vector; an anti-Kidins220 antibody (used in the preparation of Kidins220)  
CC ; an agent (A1) which modulates activity of Kidins220, protein kinase D  
CC (PKD) or interaction between PKD and Kidins220; a mutant animal  
CC transgenic for kidins220; the use of an agent capable of detecting the  
CC expression of Kidins220 gene in manufacture of a diagnostic reagent for  
CC diagnosing or prognosing cancer or for monitoring the progression of  
CC cancer in a patient; a compound comprising a moiety that selectively  
CC binds to Kidins220 or its variant and another moiety; and a nucleic acid  
CC molecule encoding the compound. The antibody is useful for modulating the  
CC activity of Kidins220, where the antibody prevents a region of Kidins220  
CC interacting from another protein, or reduces the ability of Kidins220 to  
CC bind to adenosine triphosphate (ATP). The kidins220 polynucleotide is  
CC useful for identifying an agent which modulates the promoter activity of  
CC the polynucleotide, and Kidins220 is useful for identifying an agent

CC which modulates activity of Kidins220, protein kinase D (PKD) and the  
 CC interaction between PKD with Kidins220. Kidins220 or its polynucleotide  
 CC is useful in medicine, e.g. using gene therapy or for diagnosing or  
 CC monitoring progression of cancer in a patient. The agent is useful in  
 CC manufacture of medicament, for use in treatment of neurodegenerative  
 CC disease, and the agent which inhibits function of Kidins220 gene or its  
 CC product is useful for treating cancer which is glioblastoma multiforme or  
 CC prostate cancer. The compound is useful for imaging cancer in an  
 CC individual, and for diagnosing or prognosing, and also for treating an  
 CC individual. The present sequence represents Human Kidins220  
 XX  
 SQ Sequence 1715 AA;

Query Match 100.0%; Score 8884; DB 5; Length 1715;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVLISQSVINYVEENIPALKALLECKDVDERNECGQTPLMAAEOGNLEIVKELIKN 60  
 DB 1 MSVLISQSVINYVEENIPALKALLECKDVDERNECGQTPLMAAEOGNLEIVKELIKN 60  
 QY 61 GANCNLEDLNNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGMWTALMWACYKGRDGV 120  
 DB 61 GANCNLEDLNNWTALISASKEGHVHIVEELLKCGVNLEHRDMGGMWTALMWACYKGRDGV 120  
 QY 121 ELLLSHGANSVTGLQSVYVPIIWAAGRGHADIHVLQLLQNGAKVNSDKYGTTPVVAAR 180  
 DB 121 ELLLSHGANSVTGLQSVYVPIIWAAGRGHADIHVLQLLQNGAKVNSDKYGTTPVVAAR 180  
 QY 181 KGHLECVKHLAMGADVQGGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDGNTA 240  
 DB 181 KGHLECVKHLAMGADVQGGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDGNTA 240  
 QY 241 LMTASKEGHTIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
 DB 241 LMTASKEGHTIEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
 QY 301 QDNKTALYWAVERGNATWVRDILQCPNPDTEICTKDGTEPLIKATKRNTEVVELLDKGA 360  
 DB 301 QDNKTALYWAVERGNATWVRDILQCPNPDTEICTKDGTEPLIKATKRNTEVVELLDKGA 360  
 QY 361 KVSADVKKGTPLHAIIRGRSKLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420  
 DB 361 KVSADVKKGTPLHAIIRGRSKLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420  
 QY 421 TQIFGARHLSPTEITDGMGLGYDLYSSALADILSEPTMQPPI CVGLYAQMGSGKSFLLKKL 480  
 DB 421 TQIFGARHLSPTEITDGMGLGYDLYSSALADILSEPTMQPPI CVGLYAQMGSGKSFLLKKL 480  
 QY 481 EDEMKTFAGQOIPELPQFSLIIVFLTLILCGGLGILFAFTVHNLGIAVLSFLALLYIF 540  
 DB 481 EDEMKTFAGQOIPELPQFSLIIVFLTLILCGGLGILFAFTVHNLGIAVLSFLALLYIF 540  
 QY 541 FIVIFGGRREGSGSNWAVLSRLARHIGYLELLKJLMPVNPPELPEQTTKALPVRELF 600  
 DB 541 FIVIFGGRREGSGSNWAVLSRLARHIGYLELLKJLMPVNPPELPEQTTKALPVRELF 600  
 QY 601 TDYNRLLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKTKCCLPS 660  
 DB 601 TDYNRLLSSVGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKTKCCLPS 660  
 QY 661 FVIFLFIIGCIISGITLLAIFVRDPKHLTVNAVLISIASVVGGLAFVNLNCRTHWQVLDL 720  
 DB 661 FVIFLFIIGCIISGITLLAIFVRDPKHLTVNAVLISIASVVGGLAFVNLNCRTHWQVLDL 720  
 QY 721 NSORKRLHNAASKHLKSKSGFMKVLKCEVELMARMAKTTIDSFQNTQRLVLIIDGLDAC 780  
 DB 721 NSORKRLHNAASKHLKSKSGFMKVLKCEVELMARMAKTTIDSFQNTQRLVLIIDGLDAC 780  
 QY 781 EQDKVLQMLDITVRVLSKSGFFIAFASDPHIIKAINQNLSVLRDSNNGHDMYNIHV 840  
 DB 781 EQDKVLQMLDITVRVLSKSGFFIAFASDPHIIKAINQNLSVLRDSNNGHDMYNIHV 840

QY 841 LPVELNSRGLSNARKFLVTSATNGDVPVCSDDTTGIGEDADRRVSONSLGEMTKLGSKTALN 900  
 DB 841 LPVELNSRGLSNARKFLVTSATNGDVPVCSDDTTGIGEDADRRVSONSLGEMTKLGSKTALN 900  
 QY 901 RRDYTRRRQMQRTITRQMSFDLTLLVTEDEWFSIDSPQTMRRLLNIVSVTGRLLRANQIS 960  
 DB 901 RRDYTRRRQMQRTITRQMSFDLTLLVTEDEWFSIDSPQTMRRLLNIVSVTGRLLRANQIS 960  
 QY 961 FFWDLASWINLQWPPYRTSWLLYLETEGIPDMQTLKTIYERISKNIPTTKQVPEPLL 1020  
 DB 961 FFWDLASWINLQWPPYRTSWLLYLETEGIPDMQTLKTIYERISKNIPTTKQVPEPLL 1020  
 QY 1021 EIDGDIRNEFEVFLSRTPVLVARDVKVFLPCTVNLDPKLEIADVRAAREQISIGGLAY 1080  
 DB 1021 EIDGDIRNEFEVFLSRTPVLVARDVKVFLPCTVNLDPKLEIADVRAAREQISIGGLAY 1080  
 QY 1081 PPLPLHEGPPRAPSGVQSPVCSSTSFNGPAGGVVSPQPHSSYYSYSGMTGQHPFYNRG 1140  
 DB 1081 PPLPLHEGPPRAPSGVQSPVCSSTSFNGPAGGVVSPQPHSSYYSYSGMTGQHPFYNRG 1140  
 QY 1141 SGAPGPPVLLNSLVNDAVCEKLKQIEGLDQSMLEPOYCTTIKKANINGRVLACNIDELK 1200  
 DB 1141 SGAPGPPVLLNSLVNDAVCEKLKQIEGLDQSMLEPOYCTTIKKANINGRVLACNIDELK 1200  
 QY 1201 KEMNNFQDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHT 1260  
 DB 1201 KEMNNFQDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHT 1260  
 QY 1261 ELSSTQPTTLNFSPEELNTLGLDEGAPRHSNLSWQSQTTRTPSLSLNSQDSSIEISKLT 1320  
 DB 1261 ELSSTQPTTLNFSPEELNTLGLDEGAPRHSNLSWQSQTTRTPSLSLNSQDSSIEISKLT 1320  
 QY 1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSPHSTYTMGQSSSGSIHNSLQEK 1380  
 DB 1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSPHSTYTMGQSSSGSIHNSLQEK 1380  
 QY 1381 KQSEKPDGGRKSFMLKRGDVIDYSSGVSTNDASPLDPIITEDEKSDGSKLLPGKKS 1440  
 DB 1381 KQSEKPDGGRKSFMLKRGDVIDYSSGVSTNDASPLDPIITEDEKSDGSKLLPGKKS 1440  
 QY 1441 SERSSIFQTDLKLKSGSLRYQKLPSEDESGTEESDNTPLLDKDKRKAEGKVERVPKSP 1500  
 DB 1441 SERSSIFQTDLKLKSGSLRYQKLPSEDESGTEESDNTPLLDKDKRKAEGKVERVPKSP 1500  
 QY 1501 EHSABPRTFIKAEYLSDALDKKSDSGVRSSESSPNHSLHNEVADDQLEKANLIE 1560  
 DB 1501 EHSABPRTFIKAEYLSDALDKKSDSGVRSSESSPNHSLHNEVADDQLEKANLIE 1560  
 QY 1561 LEDDSSHGKRGIPHSISGLQDPIIARMSICSEDKSKSPSECSLIASSPEENWPACQAYNL 1620  
 DB 1561 LEDDSSHGKRGIPHSISGLQDPIIARMSICSEDKSKSPSECSLIASSPEENWPACQAYNL 1620  
 QY 1621 NRTPTSTVTLNNSAPANRANQNFDMEGIRETSQVILRPSSSPNPTTIQENLKSMTK 1680  
 DB 1621 NRTPTSTVTLNNSAPANRANQNFDMEGIRETSQVILRPSSSPNPTTIQENLKSMTK 1680  
 QY 1681 SQRSSYTRLSKDPPELHAAASSESTGFGERESIL 1715  
 DB 1681 SQRSSYTRLSKDPPELHAAASSESTGFGERESIL 1715

RESULT 5

AAE32128

ID AAE32128 standard; protein; 1715 AA.

XX AAE32128;

AC AAE32128;

XX 24-MAR-2003 (first entry)

DT

XX Human cytoskeleton-associated protein, CSAP-26.

DE Human; cytoskeleton-associated protein; CSAP-26; atherosclerosis; cancer;

KW Human; cytoskeleton-associated protein; CSAP-26; atherosclerosis; cancer;

KW gene therapy.



Db 1201 KEMNIFGDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGCEPARRASHNELPHT 1260  
Qy 1261 ELSSQPTTLNPFEEELNTLGLDEGAPRHNLNWSQOTRTPSLSLNSQDSSIEISKLT 1320  
Db 1261 ELSSQPTTLNPFEEELNTLGLDEGAPRHNLNWSQOTRTPSLSLNSQDSSIEISKLT 1320  
Qy 1321 DKVQAEYRDYRYIAQMSQLEGGPGSTTISGRSSPHSTVYMGSSGGSIHNLNQEKG 1380  
Db 1321 DKVQAEYRDYRYIAQMSQLEGGPGSTTISGRSSPHSTVYMGSSGGSIHNLNQEKG 1380  
Qy 1381 KDSEKPDGKRGKFLMKRGVVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKS 1440  
Db 1381 KDSEKPDGKRGKFLMKRGVVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKS 1440  
Qy 1441 SERSSLPQTDLKLKGSGLRYQKLPSEDESGTGESNTPLKDDKDKARKGKVERVPKSP 1500  
Db 1441 SERSSLPQTDLKLKGSGLRYQKLPSEDESGTGESNTPLKDDKDKARKGKVERVPKSP 1500  
Qy 1501 EHSAEPIRTPIKAEVLSKDLKDDKSSDGSVRSSESSPHNSLHNEVADDSDLEKANLIE 1560  
Db 1501 EHSAEPIRTPIKAEVLSKDLKDDKSSDGSVRSSESSPHNSLHNEVADDSDLEKANLIE 1560  
Qy 1561 LEDDSHGKRGKGI PHSLSGLQDPIIARMSICSEDKSPSECSLIASSPEENWPACQKAYNL 1620  
Db 1561 LEDDSHGKRGKGI PHSLSGLQDPIIARMSICSEDKSPSECSLIASSPEENWPACQKAYNL 1620  
Qy 1621 NRTPTSTVTLNNGAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQENLKSMTNKR 1680  
Db 1621 NRTPTSTVTLNNGAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQENLKSMTNKR 1680  
Qy 1681 SQSSSYRLSKDPPELHAAASSESTGFGEERESIL 1715  
Db 1681 SQSSSYRLSKDPPELHAAASSESTGFGEERESIL 1715

RESULT 6

ADA09888  
ID ADA09888 standard; protein; 1753 AA.  
AC ADA09888;  
DT 20-NOV-2003 (first entry)  
XX Human receptor and membrane-associated protein REMAP-30, SEQ ID:30.  
DE Human; receptor and membrane-associated protein; REMAP;  
XX cell proliferative disorder; cancer; autoimmune disorder;  
KW inflammatory disorder; infection; neurological disorder;  
KW metabolic disorder; developmental disorder; endocrine disorder;  
KW cytostatic; immunosuppressive; antiinflammatory; neuroprotective;  
KW neurotropic; cerebroprotective; gene therapy; receptor.  
XX Homo sapiens.  
OS W02003070902-A2.  
XX PN 28-AUG-2003.  
XX PD 18-FEB-2003; 2003WO-US004902.  
XX PF 20-FEB-2002; 2002US-0358279P.  
XX PR 13-MAR-2002; 2002US-0364338P.  
XX PR 25-APR-2002; 2002US-0375657P.  
XX PR 29-APR-2002; 2002US-0376669P.  
XX PR 10-MAY-2002; 2002US-0379837P.  
XX PR 10-MAY-2002; 2002US-0379853P.  
XX (INCY-) INCYTE GENOMICS INC.  
PA Chawla NK, Yue H, Richardson TW, Marquis JP, Lehr-Mason PM;  
XX Gorvad AE, Becha SD, Kable AE, Swarnakar A, Jin P, Hawkins PR;  
PI Chien D, Ramkumar J, Tran UK, Hafalia AJA, Baughn MR, Lee SV;  
PI Jiang X, Jackson AA, Khare R, Bullock SA;

XX WPI; 2003-697610/66.  
DR N-PSDB; ADA09926.  
XX New human receptor and membrane associated proteins and nucleic acids,  
PT useful for diagnosing, treating or preventing e.g. viral, bacterial,  
PT fungal, parasitic, protozoan or helminthic infections, cancers,  
PT neurological disorders.  
XX Claim 1; Page 254-257; 298pp; English.  
PS The invention relates to 38 human receptors and membrane-associated  
XX proteins (REMAP). REMAP-1 to REMAP-38, and the cDNAs encoding them  
CC (ADA09889-ADA09934). The invention also encompasses expressing  
CC constructs, host cells and transgenic organisms comprising a REMAP  
CC nucleic acid sequence; the recombinant preparation of a REMAP; an  
CC antibody against a REMAP; methods of detection of REMAP proteins or  
CC nucleic acids; a micro-array containing REMAP nucleic acids; methods of  
CC screening compounds for their ability to modulate REMAP activity or  
CC expression; and pharmaceutical compositions comprising a REMAP protein, a  
CC REMAP agonist or REMAP antagonist. The REMAP proteins, nucleic acids or  
CC compositions comprising them are useful in diagnosing, treating or  
CC preventing a variety of disorders, including cell proliferative disorders  
CC (e.g., arteriosclerosis, cirrhosis, hepatitis, psoriasis, or primary  
CC thrombocytopenia) or cancers (e.g., adenocarcinoma, leukaemia, or cancers  
CC of the bone, brain, breast or uterus); autoimmune/inflammatory disorders  
CC (e.g., AIDS, allergies, anaemia, asthma, bronchitis, gout, multiple  
CC sclerosis, osteoarthritis, rheumatoid arthritis, or systemic lupus  
CC erythematosus); viral, bacterial, fungal, parasitic, protozoan or  
CC helminthic infections; neurological disorders (e.g., epilepsy, stroke,  
CC Alzheimer's disease, dementia, or Parkinson's disease); metabolic  
CC disorders (e.g., hereditary fructose intolerance, obesity, hypoglycaemia  
CC or diabetes); developmental disorders (e.g., achondroplastic dwarfism,  
CC hypothyroidism or hydrocephalus); or endocrine disorders (e.g., disorders  
CC of the hypothalamus or pituitary gland). The present sequence represents  
CC a REMAP of the invention.  
XX SQ Sequence 1753 AA;  
Query Match 99.7%; Score 8855; DB 6; Length 1753;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 1715; Conservative 0; Mismatches 0; Indels 38; Gaps 1;  
Qy 1 MSVLISQSVINVEENIPALKALLEKCKVDVERNECGQTPLMIAAEOGNLEIVKELIKN 60  
Db 1 MSVLISQSVINVEENIPALKALLEKCKVDVERNECGQTPLMIAAEOGNLEIVKELIKN 60  
Qy 61 GANCNLEDLNNWTALISASKEGHVHIVEELLKCGVNLHRDMGGWTALMWACYKGRTDVV 120  
Db 61 GANCNLEDLNNWTALISASKEGHVHIVEELLKCGVNLHRDMGGWTALMWACYKGRTDVV 120  
Qy 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADI VHLLLQNGAKVNCSDKYGTTPLVWAAR 180  
Db 121 ELLLSHGANSPTGLQYSVYPIIWAAGRGHADI VHLLLQNGAKVNCSDKYGTTPLVWAAR 180  
Qy 181 KGHLECVKHLANGADVDOEGANSMTALIVAKGGYTQSVKELIKENPNVNLTKDGNTA 240  
Db 181 KGHLECVKHLANGADVDOEGANSMTALIVAKGGYTQSVKELIKENPNVNLTKDGNTA 240  
Qy 241 LMIASKEGHEIVQDILLDAGTYVNI PDRSGDTVLIGAVRGHVHIVEIRALLQKYADIDIRG 300  
Db 241 LMIASKEGHEIVQDILLDAGTYVNI PDRSGDTVLIGAVRGHVHIVEIRALLQKYADIDIRG 300  
Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGGETPLIKATKVRNIEVVELLDKGA 360  
Db 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGGETPLIKATKVRNIEVVELLDKGA 360  
Qy 361 KVSADVKKGDTPLHIAIRGSRKLAELLRNPKDGLLYRPNKAGETPNIDCSHOKSIL 420  
Db 361 KVSADVKKGDTPLHIAIRGSRKLAELLRNPKDGLLYRPNKAGETPNIDCSHOKSIL 420  
Qy 421 TQIFGARHLSPTETDQDMLGYDLISALADILSEPTMQPPI CVGLYAQWNGSGKSFLLKXL 480

Db 421 TOIFGARHLSPTETDGMGLDYLSALADILSEPTMQPPICVGLYAQWGSKSLKL 480  
QY 481 EDEMTKFAQOIEPLFQFQSWLIVFTLLCGGLGALLFAFTHVHPNGLIAVSFLALLYIF 540  
Db 481 EDEMTKFAQOIEPLFQFQSWLIVFTLLCGGLGALLFAFTHVHPNGLIAVSFLALLYIF 540  
QY 541 FIVIVFGRRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPPEQTTKALPVRFLF 600  
Db 541 FIVIVFGRRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPPEQTTKALPVRFLF 600  
QY 601 TDYNRLSSVGGTSLAEMTIATLSDACERFGLATRLFRVFTKTEQCKKWKTCCLPS 660  
Db 601 TDYNRLSSVGGTSLAEMTIATLSDACERFGLATRLFRVFTKTEQCKKWKTCCLPS 660  
QY 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLSIASVUGLAFVLCNRTWQVLDL 720  
Db 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLSIASVUGLAFVLCNRTWQVLDL 720  
QY 721 NSQKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQLVVIIDGLDAC 780  
Db 721 NSQKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQLVVIIDGLDAC 780  
QY 781 EODKVLQMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLRDSNINHGYRNIIVH 840  
Db 781 EODKVLQMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLRDSNINHGYRNIIVH 840  
QY 841 LPVFLNSRGLSNARKFLVTSATNGVPCSDTTGIOEDADRRVSONSGEMTKGSKTALN 900  
Db 841 LPVFLNSRGLSNARKFLVTSATNGVPCSDTTGIOEDADRRVSONSGEMTKGSKTALN 900  
QY 901 RRDYRRRQMQRTITRQMSFDLTKLLVTEDFSDISPOQMRRLNIVSVTCRLFRANQIS 960  
Db 901 RRDYRRRQMQRTITRQMSFDLTKLLVTEDFSDISPOQMRRLNIVSVTCRLFRANQIS 960  
QY 961 FNWDLASWINLTEQWPYRTSLIILYLBETEGIPQWTLKIYERI SKNIPTTKDVEPLL 1020  
Db 961 FNWDLASWINLTEQWPYRTSLIILYLBETEGIPQWTLKIYERI SKNIPTTKDVEPLL 1020  
QY 1021 EIDGDIRNFVFLSRTFVLVARDVKVFLPCTVNDPKLREIADVRAAREQISIGGLAY 1080  
Db 1021 EIDGDIRNFVFLSRTFVLVARDVKVFLPCTVNDPKLREIADVRAAREQISIGGLAY 1080  
QY 1081 PPLPLHEGPAPSGYSPSCSTSFNGPPAGVSPQPHSSYSGMTGPQHPFYNR - 1139  
Db 1081 PPLPLHEGPAPSGYSPSCSTSFNGPPAGVSPQPHSSYSGMTGPQHPFYNR 1140  
QY 1140 -----GSGPAGPVVLLNSLNDVAVCEK 1162  
Db 1141 FPAPLYTTRYYPGSGOHLISRPVSKTSLPRDQNNGLSGPAGPVVLLNSLNDVAVCEK 1200  
QY 1163 LKQIEGLDQSMLOPYCTTIKKANINGRVLAOCNIDELKEMNFGDWHLFRSTVLEWRN 1222  
Db 1201 LKQIEGLDQSMLOPYCTTIKKANINGRVLAOCNIDELKEMNFGDWHLFRSTVLEWRN 1260  
QY 1223 AESHVVPEDPRLSSSSGAPHGPARASHNELPHTLSQTPYTLNFSFEELNTLGL 1282  
Db 1261 AESHVVPEDPRLSSSSGAPHGPARASHNELPHTLSQTPYTLNFSFEELNTLGL 1320  
QY 1283 DEGAPRHNLSQSTRTPSLSSLSNSQDSSTISKLTQVQAEYRDYRVIYIAQMSOLE 1342  
Db 1321 DEGAPRHNLSQSTRTPSLSSLSNSQDSSTISKLTQVQAEYRDYRVIYIAQMSOLE 1380  
QY 1343 GPGGTTISGRSSPHSTYMGSSSGSIHNLQEKGDSEPKDDGKRSFLMKRGDVI 1402  
Db 1381 GPGGTTISGRSSPHSTYMGSSSGSIHNLQEKGDSEPKDDGKRSFLMKRGDVI 1440  
QY 1403 DYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKSSSRSSLFQTDLKLKSGSLRYQK 1462  
Db 1441 DYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKSSSRSSLFQTDLKLKSGSLRYQK 1500  
QY 1463 LPSDEDESGTBSNDTPLLKDDKORKAEGKVERVPKSPHSAEPIRTPIKAEYLSDAL 1522  
Db 1501 LPSDEDESGTBSNDTPLLKDDKORKAEGKVERVPKSPHSAEPIRTPIKAEYLSDAL 1560

QY 1523 DKQSSDSGVSRSSSPNHSLNHVADDSQLEKANLIELEDDSHSGKRGIPHSLSGLQDP 1582  
Db 1561 DKQSSDSGVSRSSSPNHSLNHVADDSQLEKANLIELEDDSHSGKRGIPHSLSGLQDP 1620  
QY 1583 IIA RMSICSEDKKSPSECSLIASSPEENWPAQKAYNLNRTPTSTVTLNNSAPANRANQN 1642  
Db 1621 IIA RMSICSEDKKSPSECSLIASSPEENWPAQKAYNLNRTPTSTVTLNNSAPANRANQN 1680  
QY 1643 FDEMEGIRETSQVILRPSSSPNPTTQNLKSMTHKESQSSYTRLSKDPPELHAAASS 1702  
Db 1681 FDEMEGIRETSQVILRPSSSPNPTTQNLKSMTHKESQSSYTRLSKDPPELHAAASS 1740  
QY 1703 ESTGFGERESIL 1715  
Db 1741 ESTGFGERESIL 1753

RESULT 7  
ADC06847

ID ADC06847 standard; protein; 1771 AA.

XX ADC06847;

AC ADC06847;  
XX DT 18-DEC-2003 (first entry)

XX Human Kidins220Pc protein AB033076.

DE cytostatic; prostate cancer; breast; gene therapy; transgenic; human;  
XX KW Kidins220Pc; kinase D-interacting substrate of 220kDa; chromosome 2p25.1.

XX OS Homo sapiens.

XX PN WO2003064599-A2.

XX PD 07-AUG-2003.

XX PF 24-JAN-2003; 2003WO-US001943.

XX PR 25-JAN-2002; 2002US-00054935.

PR 14-FEB-2002; 2002US-0356130P.

PR 22-MAR-2002; 2002US-00102946.

PR 08-APR-2002; 2002US-00117229.

PR 14-MAY-2002; 2002US-00144198.

PR 19-JUL-2002; 2002US-00197824.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

XX Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;

XX WPI; 2003-679495/64.

XX New isolated polynucleotide related to cancer genes, useful for  
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing  
PT or treating cancers, e.g. breast and prostate cancers.

XX Disclosure; Fig 20; 128pp; English.

XX The invention relates to a novel isolated polynucleotide comprising a  
CC differentially-regulated mammalian cancer gene. The polynucleotides of  
CC the invention demonstrate cytostatic activity and are differentially  
CC expressed in prostate cancer. The polynucleotide, polypeptides and  
CC methods of the invention may be useful for detecting, diagnosing,  
CC staging, monitoring, prognosticating, preventing or treating cancers,  
CC particularly breast and prostate cancers. Furthermore, the invention may  
CC be utilised during gene therapy procedures or in the production of  
CC transgenic animals. The current sequence is that of the prostate cancer-  
CC related protein of the invention. The current sequence is not fully  
CC legible within the specification; the complete sequence was obtained from  
CC GenBank.

XX Sequence 1771 AA;

Query Match		99.4%;	Score 8830;	DB 7;	Length 1771;
Best Local Similarity		96.7%;	Pred. No. 0;		
Matches 1714;		Conservative	0;	Mismatches	0;
				Indels	58;
				Gaps	2;
Qy	1	MSVLISQSVINYVEEENIPALKALLECKVDKVDNERNECGQTPLMIAAEOGNLEIVKELIKN	60		
Db	1	MSVLISQSVINYVEEENIPALKALLECKVDKVDNERNECGQTPLMIAAEOGNLEIVKELIKN	60		
Qy	61	GANCNLEDLONWTALISASKEGHVHIVVEELLKCGVNLHRDMGGWTALMAWCKRGRTDVV	120		
Db	61	GANCNLEDLONWTALISASKEGHVHIVVEELLKCGVNLHRDMGGWTALMAWCKRGRTDVV	120		
Qy	121	ELLISHGANPVTGLQVSVPIIWAAGRHADIVHLLQLONGAKVNCSDKYGTTPLVWAAR	180		
Db	121	ELLISHGANPVTGL-YSVPIIWAAGRHADIVHLLQLONGAKVNCSDKYGTTPLVWAAR	179		
Qy	181	KGHLECVKHLAMGADVDQEGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDQNTA	240		
Db	180	KGHLECVKHLAMGADVDQEGANSMTALIIVAVKGGYTQSVKEILKRNPNVNLTKDQNTA	239		
Qy	241	LMIASKEGHTIEIVODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	300		
Db	240	LMIASKEGHTIEIVODLLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG	299		
Qy	301	QDNKTALYWAEGKGNATWVRDILQCNPDTEICTKOGETPLIKATQWNIIEVVELLDKGA	360		
Db	300	QDNKTALYWAEGKGNATWVRDILQCNPDTEICTKOGETPLIKATQWNIIEVVELLDKGA	359		
Qy	361	KVSAVDKGGTPTLHIAIRGSRKLAELLRNPKDGRLLYRPNKAGETPPYINDCSHQKSIL	420		
Db	360	KVSAVDKGGTPTLHIAIRGSRKLAELLRNPKDGRLLYRPNKAGETPPYINDCSHQKSIL	419		
Qy	421	TQIFGARHLSPETDGMGLYDLYSSALADILSEPTWQBPICVGLYAQWGSFGSLKKL	480		
Db	420	TQIFGARHLSPETDGMGLYDLYSSALADILSEPTWQBPICVGLYAQWGSFGSLKKL	479		
Qy	481	EDEMKTFAGQOIRPLFOFSLWLVFLTLILCGGLLFAFTVHNLGIAVLSFLALIYIF	540		
Db	480	EDEMKTFAGQOIRPLFOFSLWLVFLTLILCGGLLFAFTVHNLGIAVLSFLALIYIF	539		
Qy	541	FIVYFGRRREGESWNWVLSRLARHIGYLELLLLKLMFVNPPPEQTTKALPVRFLF	600		
Db	540	FIVYFGRRREGESWNWVLSRLARHIGYLELLLLKLMFVNPPPEQTTKALPVRFLF	599		
Qy	601	TDYNRLSSVGGETSLAEMIATLSDACBRFGFLATRLFRVFKTEDTGKKKKWTKCCLPS	660		
Db	600	TDYNRLSSVGGETSLAEMIATLSDACBRFGFLATRLFRVFKTEDTGKKKKWTKCCLPS	659		
Qy	661	FVIFLFIIGCTIIGITILLAIFRVDPKHLTVNAVLISIASVVGGLAFVNLNCRTWQVLDL	720		
Db	660	FVIFLFIIGCTIIGITILLAIFRVDPKHLTVNAVLISIASVVGGLAFVNLNCRTWQVLDL	719		
Qy	721	NSQRKRLHNAASKLHLKLBGFMKVLKCEVELMARMAKTIIDSFQNTQRLVLIIDGLDAC	780		
Db	720	NSQRKRLHNAASKLHLKLBGFMKVLKCEVELMARMAKTIIDSFQNTQRLVLIIDGLDAC	779		
Qy	781	EQDKVQLMDTVRVLFSKGPPIAIPASDPHIIKAINQNLSVLDRSNNINGHYMRNIVH	840		
Db	780	EQDKVQLMDTVRVLFSKGPPIAIPASDPHIIKAINQNLSVLDRSNNINGHYMRNIVH	839		
Qy	841	LPVFLNSRGLSNARKFLVTSATNGDVPVCSDTTGIQEDADRRVQNSIGEMTKLGSKTALN	900		
Db	840	LPVFLNSRGLSNARKFLVTSATNGDVPVCSDTTGIQEDADRRVQNSIGEMTKLGSKTALN	899		
Qy	901	RRDTYRRQWRTITRQMSFDLTKLLVTEDFWSDISQPMRRLLINIVSVTGRLLRANQIS	960		
Db	900	RRDTYRRQWRTITRQMSFDLTKLLVTEDFWSDISQPMRRLLINIVSVTGRLLRANQIS	959		
Qy	961	FNMDRLASWINLTQWPYRTSWLLIYLEETEGIPDQWTLKTIYVERISKNIPTTKDVEPLL	1020		
Db	960	FNMDRLASWINLTQWPYRTSWLLIYLEETEGIPDQWTLKTIYVERISKNIPTTKDVEPLL	1019		
Qy	1021	EIDGDIRNEFVLSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGGLAY	1080		

Db	1020	EIDGDIRNEFVLSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGGLAY	1079		
Qy	1081	PPLPLEGPPRAPSGYSQPPSVCSSTSPNGPAGGVVSPQPHSSYYSGMTGPQHPFYN--	1138		
Db	1080	PPLPLEGPPRAPSGYSQPPSVCSSTSPNGPAGGVVSPQPHSSYYSGMTGPQHPFYNRP	1139		
Qy	1139	-----RSGSP	1143		
Db	1140	FPAPYLYTRYYPGGSHLISRPVKTSLPRDQNGLEVIKEDAAEGLSPPTOSSRSGSP	1199		
Qy	1144	APGPVLLANSLNDVACEKLKQIEGLDQSMPLPOYCTTIKKANINGRVLQAQCNIDELKKEM	1203		
Db	1200	APGPVLLANSLNDVACEKLKQIEGLDQSMPLPOYCTTIKKANINGRVLQAQCNIDELKKEM	1259		
Qy	1204	NNMFGDWHLFRSTVLEMRNAESHVVPEDPRFISESSSGPAPGEPARRASHNELPHTELS	1263		
Db	1260	NNMFGDWHLFRSTVLEMRNAESHVVPEDPRFISESSSGPAPGEPARRASHNELPHTELS	1319		
Qy	1264	SQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSOTRTPSLSSLSNSQDSIEISKLTDKV	1323		
Db	1320	SQTPYTLNFSFEELNTLGLDEGAPRHSNLSWQSOTRTPSLSSLSNSQDSIEISKLTDKV	1379		
Qy	1324	QAEYRDAYREYIAQMSQLEGPGSTTISGRSSPHSYTYMQSSGSGSIHSNLQEKGKDS	1383		
Db	1380	QAEYRDAYREYIAQMSQLEGPGSTTISGRSSPHSYTYMQSSGSGSIHSNLQEKGKDS	1439		
Qy	1384	EPKPDGRKSFMLMKRGDVIYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKSSER	1443		
Db	1440	EPKPDGRKSFMLMKRGDVIYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKSSER	1499		
Qy	1444	SSLFOTDLKKGSLGYOKLPSDEDESGTEESDNTPLLKDDKDKAEGKVERVPKSPHS	1503		
Db	1500	SSLFOTDLKKGSLGYOKLPSDEDESGTEESDNTPLLKDDKDKAEGKVERVPKSPHS	1559		
Qy	1504	AEPIRTFIKAYELVDALLDKDSDSGVRSSESSPNHSLHNEVADDQSLKANKLIELED	1563		
Db	1560	AEPIRTFIKAYELVDALLDKDSDSGVRSSESSPNHSLHNEVADDQSLKANKLIELED	1619		
Qy	1564	DSHSGKRGIPHSLGLODPIIARMSTCSBDKXKSPSCSLIASSPENWPAQKAYNLNRT	1623		
Db	1620	DSHSGKRGIPHSLGLODPIIARMSTCSBDKXKSPSCSLIASSPENWPAQKAYNLNRT	1679		
Qy	1624	PSVTTLNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPPTTIONENLKSMTKRSQR	1683		
Db	1680	PSVTTLNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPPTTIONENLKSMTKRSQR	1739		
Qy	1684	SSYTRLSKDPPELHAAAASSESTGFGEERESIL	1715		
Db	1740	SSYTRLSKDPPELHAAAASSESTGFGEERESIL	1771		
RESULT 8					
AAU96841					
ID	AAU96841 standard; protein; 1762 AA.				
XX	AAU96841;				
XX					
DT	09-SEP-2004 (revised)				
DT	30-JUL-2002 (first entry)				
XX					
DE	Rat kidins220 protein from Genbank AAG35185.				
XX					
KW	Kidins220; kinase D interacting substrate of 22kDa; cytosstatic;				
KW	neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;				
XX	neurodegenerative disease; glioblastoma multiforme; prostate cancer.				
OS	Rattus sp.				
XX	Unidentified.				
XX	WO200220786-A2.				
XX					
PD	14-MAR-2002.				



XX 06-SEP-2001; 2001WO-GB003977.  
PF  
XX  
XX  
PR 06-SEP-2000; 2000US-0230449P.  
XX  
XX  
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX  
XX Schiavo G, Iglesias T;  
XX WPI; 2002-371879/40.  
DR  
XX  
XX Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,  
PT useful for identifying modulators useful in treating cancer and  
PT neurodegenerative diseases.  
XX  
XX Disclosure; Fig 18; 228pp; English.  
PS  
XX  
CC The invention relates to polypeptide comprising a kinase D interacting  
CC substrate of 220 kDa (Kidins220) from rat and human or their fragment,  
CC variant or fusion provided that the protein is not the polypeptide  
CC encoded by GenBank Accession No. AB033076. Also included are a  
CC polynucleotide encoding Kidins220 provided that polynucleotide is not any  
CC one of the clones corresponding to the 61 GenBank Accession Nos. given in  
CC the specification; an expression/replicable vector comprising the  
CC polynucleotide; a recombinant host cell containing the polynucleotide or  
CC vector; an anti-Kidins220 antibody (used in the preparation of Kidins220)  
CC ; an agent (A1) which modulates activity of Kidins220, protein kinase D  
CC (PKD) or interaction between PKD and Kidins220; a mutant animal  
CC transgenic for Kidins220; the use of an agent capable of detecting the  
CC expression of Kidins220 gene in manufacture of a diagnostic reagent for  
CC diagnosing or prognosing cancer or for monitoring the progression of  
CC cancer in a patient; a compound comprising a moiety that selectively  
CC binds to Kidins220 or its variant and another moiety; and a nucleic acid  
CC molecule encoding the compound. The antibody is useful for modulating the  
CC activity of Kidins220, where the antibody prevents a region of Kidins220  
CC interacting from another protein, or reduces the ability of Kidins220 to  
CC bind to adenosine triphosphate (ATP). The Kidins220 polynucleotide is  
CC useful for identifying an agent which modulates the promoter activity of  
CC the polynucleotide, and Kidins220 is useful for identifying an agent  
CC which modulates activity of Kidins220, protein kinase D (PKD) and the  
CC interaction between PKD with Kidins220. Kidins220 or its polynucleotide  
CC is useful in medicine, e.g. using gene therapy or for diagnosing or  
CC monitoring progression of cancer in a patient. The agent is useful in  
CC manufacture of medicament, for use in treatment of neurodegenerative  
CC disease, and the agent which inhibits function of Kidins220 gene or its  
CC product is useful for treating cancer which is glioblastoma multiforme or  
CC prostate cancer. The compound is useful for imaging cancer in an  
CC individual, and for diagnosing or prognosing, and also for treating an  
CC individual. The present sequence represents Rat Kidins220 protein from  
CC Genbank accession number AAG35185  
CC  
CC Revised record issued on 09-SEP-2004 : Correction to Organism field  
XX  
XX Sequence 1762 AA;  
SQ  
Query Match 94.1%; Score 8362.5; DB 5; Length 1762;  
Best Local Similarity 91.6%; Pred. No. 0;  
Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;  
Qy 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECQOTPLMTAAEQGNLEIVKELIKN 60  
Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECQOTPLMTAAEQGNLEIVKELIKN 60  
Qy 61 GANCNLEDLDNWTALISAKGHHVIEELLKCGVNLNHRDMGGWTALMWACYKGRDYY 120  
Db 61 GANCNLEDLDNWTALISAKGHHVIEELLKCGVNLNHRDMGGWTALMWACYKGRDYY 120  
Qy 121 ELLLSHGANSPTVGLQYVYPIIWAAGRGHADI VHLILLQNAKNCSDKYGTTPLVWAAR 180  
Db 121 ELLLSHGANSPTVGLQYVYPIIWAAGRGHADI VHLILLQNAKNCSDKYGTTPLVWAAR 179  
Qy 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 240  
Db 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 239

Db 180 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGNNTA 239  
Qy 241 LMIASKEGHTIEIVODLLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300  
Db 240 LMIASKEGHTIEIVODLLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 299  
Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360  
Db 300 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 359  
Qy 361 KVSADVKKGTDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRNKAGETPYNTDCSHQKIL 420  
Db 360 KVSADVKKGTDTPLHIAIRGRSRKLAELLLRNPKDGRLLYRNKAGETPYNTDCSHQKIL 419  
Qy 421 TQIFGARHLSPTETDGMVGYDLYSSALADILTSBPTWOPPICVGLYAGWGGKGFLLKKL 480  
Db 420 TQIFGARHLSPTETDGMVGYDLYSSALADILTSBPTWOPPICVGLYAGWGGKGFLLKKL 479  
Qy 481 EDEMKTFAGQIEPLFQFQSWLIVELTLLCGGLGLFAFTVHPNLGIAVSLFLALYIF 540  
Db 480 EDEMKTFAGQIEPLFQFQSWLIVELTLLCGGLGLFAFTVHPNLGIAVSLFLALYIF 539  
Qy 541 FIVYFGRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPEQTTKALPVRPLF 600  
Db 540 FIVYFGRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPEQTTKALPVRPLF 599  
Qy 601 TDYNRLSSVGETSLAEMIAATLSDACEREGFLATRLFRVFKTEDTGKKKWKTCCLPS 660  
Db 600 TDYNRLSSVGETSLAEMIAATLSDACEREGFLATRLFRVFKTEDTGKKKWKTCCLPS 659  
Qy 661 FVIFLFTIGCTITGILLAFRVDPKHLYTNAVLISIASVVGFLAVLNCRTWQVLDL 720  
Db 660 FVIFLFTIGCTITGILLAFRVDPKHLYTNAVLISIASVVGFLAVLNCRTWQVLDL 719  
Qy 721 NSQRKRLHNAASKLUHKLKSGFMKLVCEVELMARMARKTIDSTFQNTQLVVIIDGLDAC 780  
Db 720 NSQRKRLHNAASKLUHKLKSGFMKLVCEVELMARMARKTIDSTFQNTQLVVIIDGLDAC 779  
Qy 781 EQDKVQLMDTVRVLFSKGPPIALFASDPHPIIKAINQNLNSVLDSNINSHDYMRLNVH 840  
Db 780 EQDKVQLMDTVRVLFSKGPPIALFASDPHPIIKAINQNLNSVLDSNINSHDYMRLNVH 839  
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDSTTGTEADARRVSNLSGEMTKLASKTALN 900  
Db 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTEQEDTDRVSNLSGEMTKLASKTALN 899  
Qy 901 RRDYRRRQMQRTITROMSFDLTKLLVTEWDFSDISQPMRRLNIVSVTGRLLRANOIS 960  
Db 900 RRDYRRRQMQRTITROMSFDLTKLLVTEWDFSDISQPMRRLNIVSVTGRLLRANOIT 959  
Qy 961 FNDRLASWNLTEQWPYRTSWLLYLEETEGIPDMQTLTIYERISKNIPTTKDVEPLL 1020  
Db 960 FNDRLASWNLTEQWPYRTSWLLYLEETEGIPDMQTLTIYERISKNIPTTKDVEPLL 1019  
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLNDPKLREIIADVRAAREQISIGGLAY 1080  
Db 1020 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLNDPKLREIIADVRAAREQINIGGLAY 1079  
Qy 1081 PPLPLHGGPPRAGSYQPPSVCSSTSFNGPFAGGVSVSPQSHSSYYSGMTQPPHFNRP 1139  
Db 1080 PPLPLHGGPPRAGSYQPPSVCSSTSFNGPFAGGVSVSPQSHSSYYSGMTQPPHFNRP 1139  
Qy 1140 -----GSG-----PAPGPVLL 1151  
Db 1140 FPAPYLTPRYPGGSQHLISRSVKTSLPRDQNGLPDQSGFNKQKQAAVATGSSLL 1199  
Qy 1152 NSLNVDAVCBKLEQIEGLDQSMPLPYCTTIKKANINGRVLQAQCNIDELKKEMNNFGDWH 1211  
Db 1200 SSMVTVVCEKLEQIEGLDQSMPLPYCTTIKKANINGRVLQAQCNIDELKKEMNNFGDWH 1259  
Qy 1212 LFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHTLELSSQTPYTLN 1271  
Db 1260 LFRSMVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHTLELSSQTPYTLN 1319

QY 1272 FSEELNTLGLDRGAPHSNLSWQSOTRTPSPSLSSNSODSSIEISKLTDKVQAEYRDAY 1331  
Db 1320 FSEELNTLGLDRGAPHSNLSWQSOTRTPSPSLSSNSODSSIEISKLTDKVQAEYRDAY 1379  
QY 1332 REYIAQMSQLEGGPGSTTISGRSPHSTYYMGQSSSGGSIHNSLEQKKGKDSPEKPDQGR 1391  
Db 1380 REYIAQMSQLEGGPGSTTISGRSPHSTYYMGQSSSGGSIHNSLEQKKGKDSPEKPDQGR 1439  
QY 1392 KSFPMKRGDVIDYSSGVSNTDASPLDPIERDEKSDQSGSKLLPGKKSERSLFTDL 1451  
Db 1440 KSFPMKRGDVIDYSSGVSNTDASPLDPIERDEKSDQSGSKLLPGKKSERSLFTDL 1499  
QY 1452 KLGSGRLRYOKLPSDEDESGTSTEDNTPLKDKDKRKAEGKVERVPKSPHSAEPIRTFI 1511  
Db 1500 KLGSGRLRYOKLPSDEDESGTSTEDNTPLKDKDKRKAEGKVERVPKSPHSAEPIRTFI 1559  
QY 1512 KAXEYLSDALLDKSDSGVRSSESPHNSHLENEVADDSQLEKANLIELEDDSHGKRG 1571  
Db 1560 KAXEYLSDALLDKSDSGVRSSESPHNSHLENEVADDSQLEKANLIELEDDSHGKRG 1619  
QY 1572 IPHSLGLOPPIIARMSICSEDKSPSECSLIASSPEENWPACOKAYNLNRTSTVTLLN 1631  
Db 1620 MPHSLGLOPPIIARMSICSEDKSPSECSLIASSPEESWPAOKAYNLNRTSTVTLLN 1679  
QY 1632 NSAPANRANQNFEMEGIRETSOVLIRPSSSPNPTTIONENLKSMTKRSORSYTRLSK 1691  
Db 1680 NTAPTNRANQNFEMEGIRETSOVLIRPSSSPNPTTIONENLKSMTKRSORSYTRLSK 1739  
QY 1692 DPPELHAAASSESTGFGEERESIL 1715  
Db 1740 DASELH-AAASSESTGFGEERESIL 1762

## RESULT 9

ADCO6848  
ID ADCO6848 standard; protein; 1762 AA.  
XX  
AC ADCO6848;  
XX  
DT 18-DEC-2003 (first entry)  
DE Rat Kidins220Pc protein AF2339045.  
XX  
KW cytostatic; prostate cancer; breast; gene therapy; transgenic; rat;  
KW Kidins220Pc; kinase D-interacting substrate of 220KDa.  
XX  
OS Rattus sp.  
XX  
PN WQ2003064599-A2.  
XX  
PD 07-AUG-2003.  
XX  
XX 24-JAN-2003; 2003WO-US001943.  
XX  
XX 25-JAN-2002; 2002US-00054935.  
PR 14-FEB-2002; 2002US-0356130P.  
PR 22-MAR-2002; 2002US-00102946.  
PR 08-APR-2002; 2002US-00117229.  
PR 14-MAY-2002; 2002US-00144198.  
PR 19-JUL-2002; 2002US-00197824.  
XX  
XX (ORIG-) ORIGENE TECHNOLOGIES INC.  
XX  
XX Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;  
XX WPI; 2003-679495/64.  
XX  
XX New isolated polynucleotide related to cancer genes, useful for  
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing  
PT or treating cancers, e.g. breast and prostate cancers.  
XX  
XX Disclosure; Fig 20; 128pp; English.

XX The invention relates to a novel isolated polynucleotide comprising a  
CC differentially-regulated mammalian cancer gene. The polynucleotides of  
CC the invention demonstrate cytosolic activity and are differentially  
CC expressed in prostate cancer. The polynucleotide, polypeptides and  
CC methods of the invention may be useful for detecting, diagnosing,  
CC staging, monitoring, prognosticating, preventing or treating cancers,  
CC particularly breast and prostate cancers. Furthermore, the invention may  
CC be utilized during gene therapy procedures or in the production of  
CC transgenic animals. The current sequence is that of the prostate cancer-  
CC related protein of the invention. The current sequence is not fully  
CC legible within the specification; the complete sequence was obtained from  
CC GenBank.

XX Sequence 1762 AA;

QY 1 MSVLISSQSVINVEENIPALKKLEKCKVDVERNECGQTPLMIAAEQGNLEIVKELIKN 60  
Db 1 MSVLISSQSVINVEENIPALKKLEKCKVDVERNECGQTPLMIAAEQGNVEIVKELIKN 60  
QY 61 GANCNLELDNNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMACYKGRDVV 120  
Db 61 GANCNLELDNNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMACYKGRDVV 120  
QY 121 ELLLSHGANSPTVGLQYVYPIIWAAGRGHADIHLLQNGAKVNCSDKYGTTPLYWAAR 180  
Db 121 ELLLSHGANSPTVGLQYVYPIIWAAGRGHADIHLLQNGAKVNCSDKYGTTPLYWAAR 179  
QY 181 KGHLECVKHLAMGADVDQEGANSMTALIVAKGGYTQSVKEILKKNPNVNLTKDGNTA 240  
Db 180 KGHLECVKHLAMGADVDQEGANSMTALIVAKGGYTQSVKEILKKNPNVNLTKDGNTA 239  
QY 241 LMIASKEGHEITVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
Db 240 LMIASKEGHEITVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299  
QY 301 QDNKTALYWAKEGNAWVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 360  
Db 300 QDNKTALYWAKEGNAWVRDILQCNPDTEICTKDGTEPLIKATKMRNIEVVELLDKGA 359  
QY 361 KVSADVKKGDTPHIAIRGSRKLAELLRNPKDGBLLYRPNKAGETPNYIDCSHOKSIL 420  
Db 360 KVSADVKKGDTPHIAIRGSRKLAELLRNPKDGBLLYRPNKAGETPNYIDCSHOKSIL 419  
QY 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGLKLLKCL 480  
Db 420 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSGLKLLKCL 479  
QY 481 EDEMKTFAQQOIEPLFQFSLIVFLTLCCGGIGLGLFAFTVHPNLGIAVSLSLALLYIF 540  
Db 480 EDEMKTFAQQOIEPLFQFSLIVFLTLCCGGIGLGLFAFTVHPNLGIAVSLSLALLYIF 539  
QY 541 FVIVYFGRREGESMNWAWLSTRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 600  
Db 540 FVIVYFGRREGESMNWAWLSTRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 599  
QY 601 TDYNRLSSVGGETSLSAEMIATLSDACEREFGLATRLFRVFKTETDQGGKWKKTCCPLS 660  
Db 600 TDYNRLSSVGGETSLSAEMIATLSDACEREFGLATRLFRVFKTETDQGGKWKKTCCPLS 659  
QY 661 FVIFLFIICIIISGITLLAIFRVDPKHLTVNAVLISIASVGLAFVLCNCRNTMWQVLDL 720  
Db 660 FVIFLFIICIIISGITLLAIFRVDPKHLTVNAVLISIASVGLAFVLCNCRNTMWQVLDL 719  
QY 721 NSORKELHNAASKHLKSKSEGMKVLKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDAC 780  
Db 720 NSORKELHNAASKHLKSKSEGMKVLKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDAC 779  
QY 781 EQDKVLQMLDTRVRLFSKGPFAIPASDPHIIKAINQNLNSVLRDSNINGHDMYRNIVH 840

Query Match 94.1%; Score 8362.5; DB 7; Length 1762;

Best Local Similarity 91.6%; Pred. No. 0;

Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;

Db 780 EODKVLQMLDTRVLFSGPPIAFSPDPIIIKAINQNLNSVLRSNNGHDYMRNTVH 839  
Qy 841 LPVFLNSGLSNARFLVTSATNGDVPCSDTTGICQEDARRVSONSLGEMTKLGSKTALN 900  
Db 840 LPVFLNSGLSNARFLVTSATNGDITCSDTTGTQEDTDRRVSONSLGEMTKLGSKTALN 899  
Qy 901 RRDYRRRQMTITROMSFDTKLVLVTFDFSDISPMTRRLNIVSVTGLLRANQIS 960  
Db 900 RRDYRRRQMTITROMSFDTKLVLVTFDFSDISPMTRRLNIVSVTGLLRANQIT 959  
Qy 961 FNDRLASWINLTQWPYRTSWLILYLEETEGIPQMTLKIYERISKNIPPTTKDVEPLL 1020  
Db 960 FNDRLASWINLTQWPYRTSWLILYLEETEGIPQMTLKIYERISKNIPPTTKDVEPLL 1019  
Qy 1021 EIDGDIRNFVFLSRTPLVARDVKVFLPCTVNLDPKRLRIIADVRAAREQISIGGLAY 1080  
Db 1020 EIDGDIRNFVFLSRTPLVARDVKVFLPCTVNLDPKRLRIIADVRAAREQINIGGLAY 1079  
Qy 1081 PPLPLHEGPPRAPGYSQPPSVCSSTFNGPFGVSPQPHSSYSYSGMTGPQHPFYNR- 1139  
Db 1080 PPLPLHEGPPRAPGYSQPPSVCSSTFNGPFGVSPQPHSSYSYSGMTGPQHPFYNR 1139  
Qy 1140 -----GSG-----PAPGPVVLL 1151  
Db 1140 FPAPVLYTPRYPGSQHLISRSVYKTSPLRDQNGLPDSDGFKRQQAQVAPATGSSILL 1199  
Qy 1152 NSLVNDVACEKLOTEGLDQSMLOPYCTTIKKANINGRVLQACNIDELKEMNNVFGDW 1211  
Db 1200 SSMVTVVCEKRLQTEGLDQSMLOPYCTTIKKANINGRVLQACNIDELKEMNNVFGDW 1259  
Qy 1212 LFRSTVLEMRNAESHVPEPRFLSESSGPAHPGEPARRASHNELPHELSSQTPYTLN 1271  
Db 1260 LFRSMVLEMRVSVQVPEPRFLSESSGPAHPGEPARRASHNELPHELSSQTPYTLN 1319  
Qy 1272 PSFELNLTGLDEGAPRHNSLWSQOTRTPSLSLNSQDSSEIISKJTDKQVQYRDAY 1331  
Db 1320 PSFELNLTGLDEGAPRHNSLWSQOTRTPSLSLNSQDSSEIISKJTDKQVQYRDAY 1379  
Qy 1332 REYLAQMSQLEGGGPGSTTIGRSPSHSTYVWGQSSGSGSIHNSLEQEKDSEPKPDGR 1391  
Db 1380 REYLAQMSQLEGGGPGSTTIGRSPSHSTYVWGQSSGSGSIHNSLEQEKDSEPKPDGR 1439  
Qy 1392 KSFLMKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKSSERSLSFQTDL 1451  
Db 1440 KSFLMKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKSSERSLSFQTDL 1499  
Qy 1452 KLKSGGLRYQKLPDDESGTTEESDNTPLLKDDKDKAEGKVERVPKSPHSAPIRTFI 1511  
Db 1500 KLKSGGLRYQKLPDDESGTTEESDNTPLLKDDKDKAEGKVERVPKSPHSAPIRTFI 1559  
Qy 1512 KAKEVLSDALDKDSDSGVRSSESSPNHSLHNEVADDSDLEKANLIELEDSHSGKRG 1571  
Db 1560 KAKEVLSDALDKDSDSGVRSSESSPNHSLHNEVADDSDLEKANLIELEDSHSGKRG 1619  
Qy 1572 IPHSLSLGLQDPIIARMSICSDKSPSCSLIASSPEENWPACQKAYNLNRTPTVTILNN 1631  
Db 1620 MPHSLSLGLQDPIIARMSICSDKSPSCSLIASSPEENWPACQKAYNLNRTPTVTILNN 1679  
Qy 1632 NSAPANRANQNFDEWEGIRETSQVILRPSSPNPTTTONENLKMTHKRSQSSYTRLSK 1691  
Db 1680 NTAPTNRANQNFDEWEGIRETSQVILRPSSPNPTTTONENLKMTHKRSQSSYTRLSK 1739  
Qy 1692 DPPELHAAASSESTGFGEERESIL 1715  
Db 1740 DASELH-AAASESTGFGEERESIL 1762

RESULT 10  
AAU80244  
ID AAU80244 standard; protein; 1763 AA.  
XX  
AC AAU80244;

XX  
DT 09-SEP-2004 (revised)  
30-JUL-2002 (first entry)  
XX  
DE Rat kidins220 protein.  
XX  
KW Kidins220; kinase D interacting substrate of 22kDa; cytostatic;  
neuroprotective; gene therapy; protein kinase D; PKD; cancer; rat;  
neurodegenerative disease; glioblastoma multiforme; prostate cancer.  
XX  
OS Rattus sp.  
Unidentified.  
XX  
PN WO200220786-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 06-SEP-2001; 2001WO-GB003977.  
XX  
PR 06-SEP-2000; 2000US-0230449P.  
XX  
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX  
PI Schiavo G, Iglesias T;  
WPI; 2002-371879/40.  
XX  
DR N-PSDB; ABK51214.  
XX  
PT Polypeptide kinase D interacting substrate of 220 kDa and polynucleotide,  
useful for identifying modulators useful in treating cancer and  
neurodegenerative diseases.  
XX  
PS Claim 33; Fig 1; 228pp; English.  
XX  
CC The invention relates to polypeptide comprising a kinase D interacting  
substrate of 220 kDa (Kidins220) from rat and human or their fragment,  
variant or fusion provided that the protein is not the polypeptide  
encoded by GenBank Accession No. AB033076. Also included are a  
polynucleotide encoding Kidins220 provided that polynucleotide is not any  
one of the clones corresponding to the 61 GenBank Accession Nos. given in  
the specification; an expression/replicable vector comprising the  
polynucleotide; a recombinant host cell containing the polynucleotide or  
vector; an anti-Kidins220 antibody (used in the preparation of Kidins220)  
; an agent (A1) which modulates activity of Kidins220, protein kinase D  
(PKD) or interaction between PKD and Kidins220; a mutant animal  
transgenic for Kidins220; the use of an agent capable of detecting the  
expression of Kidins220 gene in manufacture of a diagnostic reagent for  
diagnosing or prognosing cancer or for monitoring the progression of  
cancer in a patient; a compound comprising a moiety that selectively  
binds to Kidins220 or its variant and another moiety; and a nucleic acid  
molecule encoding the compound. The antibody is useful for modulating the  
activity of Kidins220, where the antibody prevents a region of Kidins220  
interacting from another protein, or reduces the ability of Kidins220 to  
bind to adenosine triphosphate (ATP). The Kidins220 polynucleotide is  
useful for identifying an agent which modulates the promoter activity of  
the polynucleotide, and Kidins220 is useful for identifying an agent  
which modulates activity of Kidins220, protein kinase D (PKD) and the  
interaction between PKD with Kidins220. Kidins220 or its polynucleotide  
is useful in medicine, e.g. using gene therapy or for diagnosing or  
monitoring progression of cancer in a patient. The agent is useful in  
manufacture of medicament, for use in treatment of neurodegenerative  
disease, and the agent which inhibits function of Kidins220 gene or its  
product is useful for treating cancer which is glioblastoma multiforme or  
prostate cancer. The compound is useful for imaging cancer in an  
individual, and for diagnosing or prognosing, and also for treating an  
individual. The present sequence represents Rat Kidins220

Revised record issued on 09-SEP-2004 : Correction to Organism field  
Sequence 1763 AA;  
Query Match 93.4%; Score 8301; DB 5; Length 1763;  
Best Local Similarity 90.9%; Pred. No. 0;

Matches 1603; Conservative 47; Mismatches 64; Indels 50; Gaps 2;		
QY	1	MSVLISQSVINYVEENIPALKALLEKCKVDVDERNECGQTPLMIAAEQGNLEIVKELIKN 60
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDVDERNECGQTPLMIAAEQGNVEIVKELIKN 60
QY	61	GANCNLEDLNTWALTISASKEGHVHVIEELLKGVNLEHRDMGWTALMWACYKGRDVV 120
Db	61	GANCNLEDLNTWALTISASKEGHIHIVEELLKSGALEHRDMGWTALMWACYKGRDVV 120
QY	121	ELLISHGANPSVTGLOYSVYPIIWAAGRHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
Db	121	ELLISHGANPSVTGLOYSVYPIIWAAGRHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
QY	181	KHLECVKHLANGADVQEGANSMTALIVAVKGYTQSVKEILKNPNVNLTKDQNTA 240
Db	181	KHLECVKHLANGADVQEGANSMTALIVAVKGYTQSVKEILKNPNVNLTKDQNTA 240
QY	241	LMTASKEGHEIIVQDLDLDACTYNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
Db	241	LMTASKEGHEIIVQDLDLDACTYNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
QY	301	QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKOGETPLIKATKARNIEVVELLDKGA 360
Db	301	QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKOGETPLIKATKARNIEVVELLDKGA 360
QY	361	KVSAVDKKGDTPLHIAIRGSRKLAELLRNPKDGLLYRPNKAGETPNYIDCSHOKSIL 420
Db	361	KVSAVDKKGDTPLHIAIRGSRRLAELLRNPKDGLLYRPNKAGETPNYIDCSHOKSIL 420
QY	421	TOIFGARHLSPTDGDMLGYDIYSSALADILSEPTMOPPICVGLYAQMGSGSFLLKKL 480
Db	421	TOIFGARHLSPTDGDMLGYDIYSSALADILSEPTMOPPICVGLYAQMGSGSFLLKKL 480
QY	481	EDEMKTFAGQIQIPLFOFSLWLVFLTLCCGGLLFAFTVHPNGLIAVSLSLALYIF 540
Db	481	EDEMKTFAGQIQIPLFOFSLWLVFLTLCCGGLVFAFTVDTNLAIASLSFLALYIF 540
QY	541	FIVIFYGRRREGSWMWVLSRLARHIGYLBELLKLMFVNPELPEOTTKALPVRF 600
Db	541	FIVIFYGRRREGSWMWVLSRLARHIGYLBELLKLMFVNPELPELAEQTTKALPVRF 600
QY	601	TDYRNLSSVCGETSLEAMITLSDACEREGFLATRLFRVFKTDTQKKKKWKTCCPLS 660
Db	601	TDYRNLSSVCGETSLEAMITLSDACEREGFLATRLFRVLRTEESQGGKKWKTCCPLS 660
QY	661	FVIFLFTIGCIIIGITITLAIIFRVDPKHLTVNAVLISIASVVGFLAVLNCRTWQVLDL 720
Db	661	FVIFLFTIGCIIIGITITLAIIFRVDPKHLTVNAVLISIASVVGFLAVLNCRTWQVLDL 720
QY	721	NSQKRLHNAASKHLKSGEGMKVLKCEVELMARMAKTIDSTQNTQRLVVIIDGLDAC 780
Db	721	NSQKRLHNAASKHLKSGEGMKVLKCEVELMARMAKTIDSTQNTQRLVVIIDGLDAC 780
QY	781	EQDKVQLMQLTVRLVLSKGPFIAPASDPHIIKAINQNLSVLRDSNNGHDMYNIHV 840
Db	781	EQDKVQLMQLTVRLVLSKGPFIAPASDPHIIKAINQNLSVLRDSNNGHDMYNIHV 840
QY	841	LPVFLNSRGLSNARKFLVTSATNGDVPSCDTTGIQEDADRRVQNSIGEMTKLGSKTALN 900
Db	841	LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGTQEDTDRRVQNSIGEMTKLGSKTALN 900
QY	901	RRTYRRRQMQRTITROMSPDLTKLVTEWDFSDISQPNWRRLLNIVSVTGRLLRANQIS 960
Db	901	RRTYRRRQMQRTITROMSPDLTKLVTEWDFSDISQPNWRRLLNIVSVTGRLLRANQIT 960
QY	961	FNNDRLASWNLTEQPYRTSWLILYEETEGIPDQWTLKTIYERISKNPITTKDVEPLL 1020
Db	961	FNNDRLASWNLTEQPYRTSWLILYEETEGIPDQWTLKTIYERISKNPITTKDVEPLL 1020
QY	1021	EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGLAY 1080
Db	1021	EIDGDIRNFVFLSSRTPVLVARDVKTFLEPCTVNLDPKLEIITADVRAAREQINIGLAY 1080

QY	1081	PPLPLHEGPPRAPSGYSQPPSVCSSTSNFGPFAGGVVSPQPHSSYYSGMTGPOHPFYNR- 1139
Db	1081	PPLPLHEGPPRPSPGYSQPPSVCSSTSNFGPFAGGVVSPQPHSSYYSGLSGQHPFYNRP 1140
QY	1140	-----GSGPAPGPVLL 1151
Db	1141	PFAPYLYTRYPPGSGQHLSRSSVKASLPRDNNGCLPCDSGPNKQORGRPCHROITVT 1200
QY	1152	NSLNVDAVCEKLEQLQIEGLDQSMLEPOYCTTIKKANINGRVLACNIDELKEMMNFDMH 1211
Db	1201	EFNDRCCRCEKLEQLQIEGLDQSMLEPOYCTTIKKANINGRVLSCNIDELKEMMNFDMH 1260
QY	1212	LFRSTVLEWRNAESHVVPEDPRFLSESSSGPAPHGEPARRASHNELPHTLSQOTPYTLN 1271
Db	1261	LFRSMVLEMRVESQVVPEDPRFLNENSAPVPHGESARRPSHTELPHTLSQOTPYTLN 1320
QY	1272	FSFEELNTLGLDEGAPRHSNLSQWQOTRTPSLSSLSNDSQSSIEISKLDKVQAEYRDAY 1331
Db	1321	FSFEELNTLGLDEGAPRHSNLSQWQOTRTPSLSSLSNDSQSSIEISKLDKVQAEYRDAY 1380
QY	1332	REYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSGSIHNSLEQEKOKGSEPKPDGR 1391
Db	1381	REYIAQMSQLEGGTGSSTISGRSSPHSTYYIGQSSGSIHSTLEQERKEGELKQEDGR 1440
QY	1392	KSFLMKRGDVIDYSSSGVSTNDASPLDPITEDEKSDQSGSKLLPGKKSSERSLFTDL 1451
Db	1441	KSFLMKRGDVIDYSSSGVSTNEASPLDPITEDEKSDQSGSKLLPGKKSSERSLFTDL 1500
QY	1452	KLKGSGLRYQKLPSDEDESGTEESDNTPLLKDDKDKAEGKVERVPKSPHSAEPIRTFI 1511
Db	1501	KLKGGGLRYQKLPSDEDESGTEESDNTPLLKDDKDKAEGKVERVCKSPHSAEPIRTFI 1560
QY	1512	KAKEYLDALLDKQSSDSGVRSSESSPNHSLHNEVADDSQLEKANLILEDDSHSGKRG 1571
Db	1561	KAKEYLDALLDKQSSDSGVRSSESSPNHSLHNEAADDQLEKANLILEDEGHSGKRG 1620
QY	1572	IPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEENWPACOKAYNLNRTPTVTLLN 1631
Db	1621	MPHSLSGLODPIIARMSICSEDKSPSECSLIASSPEESWPACOKAYNLNRTPTVTLLN 1680
QY	1632	NSAPANRANQNDMEGIRETSQVILRPSSSNPTTIQENLKSMTHKESQSSSYTRLK 1691
Db	1681	NTAPTNRANQNDMEGIRETSQVILRPSPSNPTAVQNLKSMHAKKSQSSSYTRLK 1740
QY	1692	DPPELHAAASSESTGFEERESIL 1715
Db	1741	DASELH-AAASESTGFEERESIL 1763
RESULT 11		
AAE25143		
ID	AAE25143	standard; protein; 1715 AA.
XX	AAE25143;	
AC	AAE25143;	
XX	DT	30-OCT-2002 (first entry)
XX	DE	Rat ARMS protein.
XX	KW	Ankyrin repeat-rich membrane spanning protein; ARMS; neurotrophin;
XX	KW	ephrin; receptor tyrosine kinase; growth cone; neuron; neuronal cell;
XX	KW	diagnostic; imaging; rat.
OS	Rattus norvegicus.	
XX	Key	Location/Qualifiers
FT	Region	17..390
FT	FT	/note= "Ankyrin repeat"
FT	Region	354..493
FT	FT	/note= "N-terminal region"
FT	Domain	496..518
FT	FT	/note= "Transmembrane domain"

FT	Domain	525..546
FT	/note= "Transmembrane domain"	
FT	Region	573..638
FT	/note= "Loop between TM2 and TM3"	
FT	Domain	661..680
FT	/note= "Transmembrane domain"	
FT	Domain	688..710
FT	/note= "Transmembrane domain"	
FT	Region	749..854
FT	/note= "Carboxy tail"	
FT	Region	940..1060
FT	/note= "Carboxy tail"	
FT	Region	1081..1093
FT	/note= "Polyproline stretch"	
FT	Domain	1151..1221
FT	/note= "SAM domain"	
FT	Binding-site	1713..1715
FT	/note= "PDZ binding motif"	
FN	WO200250273-A2.	
XX	27-JUN-2002.	
XX	19-DEC-2001; 2001WO-US048603.	
XX	21-DEC-2000; 2000US-0256909P.	
PR	(UWNY ) UNIV NEW YORK STATE.	
PA	Chao MV, Kong H;	
PI	WPI; 2002-508800/54.	
DR	N-PSDB; AAD41035.	
DR	Isolated ankryrin repeat-rich membrane spanning (ARMS) polypeptide that is	
PT	a target for phosphorylation by neurotrophin and ephrin receptor tyrosine	
PT	kinases, useful as a marker for growth cones.	
PS	Claim 1; Page 87-94; 136pp; English.	
PS	The invention relates to ankryrin repeat-rich membrane spanning (ARMS)	
CC	protein which is a down stream target of neurotrophin and ephrin receptor	
CC	tyrosine kinases. ARMS DNA is useful for visualising the growth cone of	
CC	neurons. ARMS protein is useful as an indicator of the biological	
CC	activity of neurotrophins and ephrins, as a marker for neuronal cells	
CC	which have the ability to undergo continued synaptic changes through	
CC	adult life or for the presence and distribution of ARMS in such neuronal	
CC	cells, or as a marker for growth cones. ARMS DNA is useful in diagnostic	
CC	and imaging methods. The present sequence is rat ARMS protein	
XX	Sequence 1715 AA;	
SQ	Query Match	93.0%; Score 8266; DB 5; Length 1715;
	Best Local Similarity	92.7%; Pred. No. 0;
	Matches 1591; Conservative	53; Mismatches 70; Indels 2; Gaps 2;
QY	1 MSVLISQSIVNYYEENIPALKALLECKDVERNECGOTPLMIAABOGNLEIIVKELTKN	60
DB	1 MSVLISQSIVNYYEENIPALKALLECKDVERNECGOTPLMIAABOGNLEIIVKELTKN	60
QY	61 GANCNLEDLNMTPALISASKEGHVHIVEALLKCGVNLEHRDMGGWTALMWACYKGRTDVV	120
DB	61 GANCNLEDLNMTPALISASKEGHVHIVEALLKCGVNLEHRDMGGWTALMWACYKGRTDVV	120
QY	121 ELLLSHGPNSTVGLOYSVYPIIWAAGRHADIIVHLILONGAKVNCSDKYGTTPLVAAAR	180
DB	121 ELLLSHGPNSTVGLOYSVYPIIWAAGRHADIIVHLILONGAKVNCSDKYGTTPLVAAAR	180
QY	181 KGLECVKHLAMGADVDEGANMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA	240
DB	181 KGLECVKHLAMGADVDEGANMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA	240
QY	241 LMIASKEGHTIEVDLLDAGTYVNIIDRSGDTVLIIGAVRGSHVEIVRALLQKYADIDIRG	300





```
QY 901 RDTYRRQMQRTTIQMSFDTKLKLVTEDFMSDISPQTMRLRLNIVSVTGRLLRANQIS 960
Db 901 RDTYRRQMQRTTIQMSFDTKLKLVTEDFMSDISPQTMRLRLNIVSVTGRLLRANQIT 960
QY 961 FNDWELASWINLTCQWYRTSWLLIYLBETEGIPQWMTLKTYIERISKNIPTTKDVEPLL 1020
Db 961 FNDWELASWINLTCQWYRTSWLLIYLBETEGIPQWMTLKTYIERISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIRNPFVLSRTPLVARDVKVFLPCTVNLDPKRLREIADVRAAREQISIGGLAY 1080
Db 1021 EIDGDIRNPFVLSRTPLVARDVKVFLPCTVNLDPKRLREIADVRAAREQINIGGLAY 1080
QY 1081 PPLPLHEGPPRAPSGYSPVCSSTSFNGPAGVSPQPHSSYSGMTQPHFPYNRG 1140
Db 1081 PPLPLHEGPPRAPSGYSPVCSSTSFNGPAGVSPQPHSSYSGMTQPHFPYNRG 1140
QY 1141 SGPARCPVVLNLSNVADVACEKLOIEGLDQSMLEPOYCTTIKKANINGRVLACQNDIDELK 1200
Db 1141 AVFATGSSLLSSMTVDVCEKLRQIEGLDQNMPOYCTTIKKANINGRVLSCQNDIDELK 1200
QY 1201 KEMNNFQDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHCEPARRASHNELPHT 1260
Db 1201 KEMANNFQDWHLFRSWLEMRNESHVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLT 1260
QY 1261 ELSQTPYTLNPFELNLTGLDEGAPRHSNLSWQSTRTTSLSSLSQDSSIEISKLT 1320
Db 1261 ELSQTPYTLNPFELNLTGLDEGAPRHSNLSWQSTRTTSLSSLSQDSSIEISKLT 1320
QY 1321 DKVQBYRDAYREYIAQMSQLEGGPGSTISGRSSPHSTYMGQSSGSGSIHSLNLEQK 1380
Db 1321 DKVQBYRDAYREYIAQMSQLEGGPGSTISGRSSPHSTYMGQSSGSGSIHSLNLEQK 1380
QY 1381 KDSBPKPDGGRKSFLMKRGDVIDYSSSGVSTNDASPLDPIITEDESKSQSGSKLLPGKKS 1440
Db 1381 KEGELKQEDGRKSFLMKRGDVIDYSSSGVSTNEASPLDPIITEDESKSQSGSKLLPGKKS 1440
QY 1441 SERSLFQDLDLKGSLRYOKLPSEDESGTEESDNTPL-LKODKORKAEKQVERVPKS 1499
Db 1441 SERPSLFQDLDLKGGLRYOKLPSEDESGTGRVOITPHCSKMIRTKRLKAKQRECAASP 1500
QY 1500 PEHSAPRTTIKAEYLSDDLKDDSGVRSSESSPHSLNINEVADDSOLEKANLI 1559
Db 1501 QEHSAPRTTIKAEYLSDDLKDDSGVRSSESSPHSLNINEVADDSOLEKANLI 1560
QY 1560 ELEDHSGKRGIPHSLSGLQDPIIARMSICSEDKSPSECSLIASSPEENWPAQOKAYN 1619
Db 1561 ELEDHSGKRGIPHSLSGLQDPIIARMSICSEDKSPSECSLIASSPEESWPAQOKAYN 1620
QY 1620 LNRTPSTVTLNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQENLKSMTHK 1679
Db 1621 LNRTPSTVTLNNTAPTNRANQNFDEIEGIRETSQVILRPSPNPTAVQENLKSMAHK 1680
QY 1680 RSQRSSYRLSKDPPDELHAASSESTGFEERESIL 1715
Db 1681 RSQRSSYRLSKDSELH-AAASSESTGFEERESIL 1715

RESULT 13
ADC06850
ID ADC06850 standard; protein; 1142 AA.
XX
AC ADC06850;
XX
DT 18-DEC-2003 (first entry)
XX
DE Kids 220Pc-related protein Pc473.
XX
KW cytosolic; prostate cancer; breast; gene therapy; transgenic;
KW Kids220Pc; kinase D-interacting substrate of 220KDa.
XX
OS Unidentified.
XX
PN WO2003064599-A2.
```

```
XX 07-AUG-2003.
PD 24-JAN-2003; 2003WO-US001943.
XX
PF 25-JAN-2002; 2002US-00054935.
XX
PR 14-FEB-2002; 2002US-0356130P.
PR 22-MAR-2002; 2002US-00102946.
PR 08-APR-2002; 2002US-00117229.
PR 14-MAY-2002; 2002US-00144198.
PR 19-JUL-2002; 2002US-00197824.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;
XX
XX WPI; 2003-679495/64.
XX
DR New isolated polynucleotide related to cancer genes, useful for
XX detecting, diagnosing, staging, monitoring, prognosticating, preventing
XX or treating cancers, e.g. breast and prostate cancers.
XX
PS Disclosure; Fig 20; 128pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising a
CC differentially-regulated mammalian cancer gene. The polynucleotides of
CC the invention demonstrate cytostatic activity and are differentially
CC expressed in prostate cancer. The polynucleotide, polypeptides and
CC methods of the invention may be useful for detecting, diagnosing,
CC staging, monitoring, prognosticating, preventing or treating cancers,
CC particularly breast and prostate cancers. Furthermore, the invention may
CC be utilised during gene therapy procedures or in the production of
CC transgenic animals. The current sequence is that of the prostate cancer-
CC related protein of the invention. The C-terminus of the current sequence
CC is illegible within the specification and as a result, residues 1143-1715
CC are absent.
XX
SQ Sequence 1142 AA;
XX
Query Match 66.8%; Score 5931; DB 7; Length 1142;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSVLSISVINVYEEENIPALKLEKCKVDNERNECQTPLMIAAEOGNLEIVKELIKN 60
Db 1 MSVLSISVINVYEEENIPALKLEKCKVDNERNECQTPLMIAAEOGNLEIVKELIKN 60
QY 61 GANCNLEDLONWTALISASKEGHVHIIVEELKCGVNLHHRDMGWTALMWACYKGRTDVV 120
Db 61 GANCNLEDLONWTALISASKEGHVHIIVEELKCGVNLHHRDMGWTALMWACYKGRTDVV 120
QY 121 ELLLSHGANSVTLQYSVYPIIWAAGRGHADIIVHLLONGAKVNCSDKYTTPLVWAAR 180
Db 121 ELLLSHGANSVTLQYSVYPIIWAAGRGHADIIVHLLONGAKVNCSDKYTTPLVWAAR 180
QY 181 KGHLECVKHLAMGADVDOEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGNTA 240
Db 181 KGHLECVKHLAMGADVDOEGANSMTALIIVAVKGYTQSVKEILKRNPNVNLTKDGNTA 240
QY 241 LMIAKESGHTIEIVQDLLDAGTYVNI PDRSGDTVLIGAVRGGHVEIVRALLQYADIDIRG 300
Db 241 LMIAKESGHTIEIVQDLLDAGTYVNI PDRSGDTVLIGAVRGGHVEIVRALLQYADIDIRG 300
QY 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGTEPLIKATMRNIEVVELLDKGA 360
Db 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKGTEPLIKATMRNIEVVELLDKGA 360
QY 361 KVSADVKKGDTPLHIAIRGRSRKLAELLRNPKDGRLLYRPNKAGETPNYDCSHQSKIL 420
Db 361 KVSADVKKGDTPLHIAIRGRSRKLAELLRNPKDGRLLYRPNKAGETPNYDCSHQSKIL 420
QY 421 TQIFGARHLSPTETDGMVGYDLYSSALADILSEFTMQPPICVGLYAQWGSKGFLLKKL 480
Db 421 TQIFGARHLSPTETDGMVGYDLYSSALADILSEFTMQPPICVGLYAQWGSKGFLLKKL 480
```

Db 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQMGSFKLLKKL 480  
Qy 481 EDEMKTFAGQOIEPLFQFSLIIVFLTLCCGGLGGLFAFTVHNLGIAVSLSFALLIYIF 540  
Db 481 EDEMKTFAGQOIEPLFQFSLIIVFLTLCCGGLGGLFAFTVHNLGIAVSLSFALLIYIF 540  
Qy 541 FIVYFGREGESWNWAWLSTRLARHIGLYLELLKLMFVNPPPELPEQTTKALPVRFLF 600  
Db 541 FIVYFGREGESWNWAWLSTRLARHIGLYLELLKLMFVNPPPELPEQTTKALPVRFLF 600  
Qy 601 TDYNRSSVGGTSLAEMIATLSDACEREFGLATRLFRVFKTEDTGKKKKWTKCCLPS 660  
Db 601 TDYNRSSVGGTSLAEMIATLSDACEREFGLATRLFRVFKTEDTGKKKKWTKCCLPS 660  
Qy 661 FVIFLFIIGCIIGITLLAIFRVDPKLITVNAVLIASVGLAFVLCNRTWQVLDL 720  
Db 661 FVIFLFIIGCIIGITLLAIFRVDPKLITVNAVLIASVGLAFVLCNRTWQVLDL 720  
Qy 721 NSOKRHLHNAASKLHLKSGFMKVLKCEVELMARMAKTIDSTFQNTQRLVWIIDGLDAC 780  
Db 721 NSOKRHLHNAASKLHLKSGFMKVLKCEVELMARMAKTIDSTFQNTQRLVWIIDGLDAC 780  
Qy 781 EQDKVLQMLDTRVFLSKGPFIAIFASDPHIIKAINQNLNSVLRDSNININGHDYMRNVH 840  
Db 781 EQDKVLQMLDTRVFLSKGPFIAIFASDPHIIKAINQNLNSVLRDSNININGHDYMRNVH 840  
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDVPDSTTGIOEDADRRVQNSIGENTKLGSKTALN 900  
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPDSTTGIOEDADRRVQNSIGENTKLGSKTALN 900  
Qy 901 RRDYRRRQORTITROMSFDLTKLVTEWDFSDISPTQWRRLLNIVSVTGRLLRANQIS 960  
Db 901 RRDYRRRQORTITROMSFDLTKLVTEWDFSDISPTQWRRLLNIVSVTGRLLRANQIS 960  
Qy 961 FNMRLASWINLQEQPYRTSWLILYEETEGIPDQWTLKTIYERISKNIPTTKDVEPLL 1020  
Db 961 FNMRLASWINLQEQPYRTSWLILYEETEGIPDQWTLKTIYERISKNIPTTKDVEPLL 1020  
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAEQISIGGLAY 1080  
Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAEQISIGGLAY 1080  
Qy 1081 PPPLHGGPPRAPSGVSPQPSVCSSTSFNGPFAGGVVSPQPHSSYSGMTGPQHPFVNRG 1140  
Db 1081 PPPLHGGPPRAPSGVSPQPSVCSSTSFNGPFAGGVVSPQPHSSYSGMTGPQHPFVNRG 1140  
Qy 1141 SG 1142  
Db 1141 SG 1142

## RESULT 14

ADCO6846  
ID ADCO6846 standard; protein; 1184 AA.

XX AC ADCO6846;

XX DT 18-DEC-2003 (first entry)

XX DE Human Kidins220Pc protein XM\_045362.

XX KW cytosolic; prostate cancer; breast; gene therapy; transgenic; human;  
KW Kidins220Pc; kinase D-interacting substrate of 220kDa; chromosome 2p25.1.  
XX OS Homo sapiens.

XX PN W02003064599-A2.

XX PD 07-AUG-2003.

XX PF 24-JAN-2003; 2003WO-US001943.

XX PR 25-JAN-2002; 2002US-00054935.

PR 14-FEB-2002; 2002US-0356130P.  
PR 22-MAR-2002; 2002US-00102946.  
PR 08-APR-2002; 2002US-00117229.  
PR 14-MAY-2002; 2002US-00144198.  
PR 19-JUL-2002; 2002US-00197824.  
XX (ORIG-) ORIGENE TECHNOLOGIES INC.  
XX PA Sun Z, Li X, Jay G, Kovacs KF, Fan W, Shu Y;  
XX PI MPI; 2003-679495/64.  
XX PT New isolated polynucleotide related to cancer genes, useful for  
PT detecting, diagnosing, staging, monitoring, prognosticating, preventing  
PT or treating cancers, e.g. breast and prostate cancers.  
XX PS Disclosure; Fig 20; 128pp; English.  
XX CC The invention relates to a novel isolated polynucleotide comprising a  
CC differentially-regulated mammalian cancer gene. The polynucleotides of  
CC the invention demonstrate cytostatic activity and are differentially  
CC expressed in prostate cancer. The polynucleotide, polypeptides and  
CC methods of the invention may be useful for detecting, diagnosing,  
CC staging, monitoring, prognosticating, preventing or treating cancers,  
CC particularly breast and prostate cancers. Furthermore, the invention may  
CC be utilised during gene therapy procedures or in the production of  
CC transgenic animals. The current sequence is that of the prostate cancer-  
CC related protein of the invention. The current sequence is not fully  
CC legible within the specification; the complete sequence was obtained from  
CC GenBank.  
XX QY Sequence 1184 AA;

Query Match 66.4%; Score 5899.5; DB 7; Length 1184;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1138; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MSVLISQSVINVEENIPALKALLECKDQVDERNECGQTPLMIAAEOGNLEIVKELIKN 60

Db 1 MSVLISQSVINVEENIPALKALLECKDQVDERNECGQTPLMIAAEOGNLEIVKELIKN 60

Qy 61 GANCNLEDLNDNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMWACYKGRDVV 120

Db 61 GANCNLEDLNDNTALISASKEGHVHIVEELLKGVNLEHRDMGWTALMWACYKGRDVV 120

Qy 121 ELLLSHGANSPTGLQYSVYPIIWAAGRCHADIVHLLQNGAKVNSDKYGTTPLVWAAR 180

Db 121 ELLLSHGANSPTGLQYSVYPIIWAAGRCHADIVHLLQNGAKVNSDKYGTTPLVWAAR 180

Qy 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGN 240

Db 181 KGHLECVKHLAMGADVDOEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTKDGN 240

Qy 241 LMIASKEGHEIIVQDILLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKYADIRG 300

Db 241 LMIASKEGHEIIVQDILLDAGTYVNIIPDRSGDTVLICAVRGHVEIVRALLQKYADIRG 300

Qy 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATKVRNIEVVELLDKGA 360

Db 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGTEPLIKATKVRNIEVVELLDKGA 360

Qy 361 KUSAVDKKGDTPHIAIRGRSRKLAELLARNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420

Db 361 KUSAVDKKGDTPHIAIRGRSRKLAELLARNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420

Qy 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQMGSFKLLKKL 480

Db 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAQMGSFKLLKKL 480

Qy 481 EDEMKTFAGQOIEPLFQFSLIIVFLTLCCGGLGGLFAFTVHNLGIAVSLSFALLIYIF 540

Db 481 EDEMKTFAGQOIEPLFQFSLIIVFLTLCCGGLGGLFAFTVHNLGIAVSLSFALLIYIF 540



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:18:20 ; Search time 40.5 Seconds  
(without alignments)  
3161.065 Million cell updates/sec

Title: US-10-021-571-4  
Perfect score: 8884  
Sequence: 1 MSVLISQSVINVEENIPA.....LHAAASSESTGFEERSIL 1715

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	557.5	6.3	1883	4	US-09-949-016-9010
2	557.5	6.3	1883	4	US-09-949-016-9011
3	557.5	6.3	1883	4	US-09-949-016-9012
4	557.5	6.3	1883	4	US-09-949-016-9013
5	557.5	6.3	1883	4	US-09-949-016-9014
6	557.5	6.3	1883	4	US-09-949-016-9015
7	557.5	6.3	1883	4	US-09-949-016-9016
8	557.5	6.3	1883	4	US-09-949-016-9017
9	555.5	6.3	1856	4	US-09-949-016-6964
10	555.5	6.3	1880	4	US-09-949-016-5876
11	555.5	6.3	1881	4	US-09-949-016-6965
12	549	6.2	2753	4	US-09-949-016-7659
13	549	6.2	2753	4	US-09-949-016-7660
14	539	6.1	3924	4	US-09-949-016-1246
15	530.5	6.0	1839	2	US-09-172-977-4
16	530.5	6.0	1839	4	US-09-404-108-4
17	522	5.9	1719	4	US-09-949-016-6966
18	520	5.9	4377	4	US-09-949-016-6978
19	493.5	5.6	843	2	US-09-172-977-3
20	493.5	5.6	843	4	US-09-404-108-3
21	480.5	5.4	1745	2	US-09-031-485-33
22	480.5	5.4	1745	2	US-08-847-429A-33
23	480.5	5.4	1745	3	US-09-065-474-33
24	480.5	5.4	1745	3	US-09-557-034-33
25	444	5.0	994	4	US-10-164-595-38
26	443.5	5.0	3913	4	US-09-949-016-10933
27	438	4.9	1088	3	US-09-082-059-2

28	437	4.9	1881	4	US-09-949-016-8307	Sequence 8307, Ap
29	409.5	4.6	683	4	US-10-164-595-71	Sequence 71, Appl
30	409.5	4.6	1423	3	US-08-810-712-10	Sequence 10, Appl
31	409.5	4.6	1431	4	US-09-538-092-1198	Sequence 1198, Ap
32	384.5	4.3	1704	4	US-09-392-812A-2	Sequence 2, Appl
33	376.5	4.2	1327	3	US-09-196-387-2	Sequence 2, Appl
34	376.5	4.2	1327	4	US-09-841-835-2	Sequence 2, Appl
35	376.5	4.2	1327	4	US-09-972-115A-8	Sequence 8, Appl
36	376.5	4.2	1619	4	US-09-392-812A-4	Sequence 4, Appl
37	374	4.2	348	2	US-09-031-485-28	Sequence 28, Appl
38	374	4.2	348	2	US-08-847-429A-28	Sequence 28, Appl
39	374	4.2	348	3	US-09-065-474-28	Sequence 28, Appl
40	374	4.2	348	3	US-09-557-034-28	Sequence 28, Appl
41	373.5	4.2	546	4	US-09-949-016-7169	Sequence 10, Appl
42	370.5	4.2	949	3	US-09-196-387-10	Sequence 10, Appl
43	370.5	4.2	949	4	US-09-841-835-10	Sequence 10, Appl
44	370	4.2	1166	4	US-09-972-115A-6	Sequence 6, Appl
45	370	4.2	1227	4	US-09-849-602-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-9010  
; Sequence 9010, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9010  
; LENGTH: 1883  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9010

Query Match	6.3%	Score	557.5	DB	4	Length	1883
Best Local Similarity	19.7%	Pred. No.	1.9e-38				
Matches	374	Conservative	286	Mismatches	659	Indels	581
Gaps	72						
Qy	21	LKALLEKCDVDERNECGQTPMIAEQNLIVKELIKNGANCNLEDDLNTALISASK	80				
Db	97	VRELNVYGANVNAQSQKGFTELYMAAQENHLEVYKFLLENGANQNVATEDGFTFLVALQ	156				
Qy	81	EGHVHIVEELLKCGV-----NLEHRDMGCGTALMWA	111				
Db	157	QGHENVVAHINVGTGKGVLPALHIAARNDDTFTAALLQNDPNDVLSKGTFTPLHIA	216				
Qy	112	CYKGRDQVVELLSHGANSPT-----GLQY-----SV	139				
Db	217	AHYENLVNAQLLNRRGASVNFPTQGITPLHIASSRRGNVIMVRLLDGRGAQIETKTDEL	276				
Qy	140	YPIWAAGRHADIVHLLONGAKVNSDKYGTTPVWAARKHLECVKHLANGADV-	198				
Db	277	TPHCAARNGHVRISIELLDHGAPIOAKTNGLSPIHMAAQGDHLCVRLQLQDAEIDD	336				
Qy	199	-----QEGA-----NSMTALIVAVKGGYTSQVKELKR	226				
Db	337	ITLDHLTPHVAHCHGHHRVAKVLLDKGAKNSPALNGFTPLHIAACKNHRVVMELLKT	396				
Qy	227	NPVNLTDKDGNTALMIASKEGTEIVQDLLDAGTYVNIIDRSQDITVLIGAVRGHVEIV	286				

Db 397 GASIDAVTESGLTPLHVASPMGHLPIVKNLQAGSPNVSNVKTPLHMAARAGTEVA 456  
QY 287 RALLQYADIDIRQDNKTALYWAKEGNATWVRDILQCNPDTEICTDQGETPLIKATKM 346  
Db 457 KYLLQNKAKVNAKADQTPHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAARE 516  
QY 347 RNTVEVELLDKAKVASVDKGDTPHIAIORSRKLAEALLLNRNPKDGRLLYRPNKAGE 406  
Db 517 GHVETVLLALEKEASOACTKGTGFTPLHVAKYGKVRVAELLE--RDA----HPNAAGK 570  
QY 407 ---TPYNDICSHQKSLITQFGARHLSPETDGMGLY--DLYSSALADILSEPTMPPIC 462  
Db 571 NGUTPLHVAHNHNLIDVKKLLPRGSPHSPAWN--GYTPLHIAKON-----QVEVA 621  
QY 463 VGLYAQWGSKSPLLKLEDEMKTFAQQTEPIQFQSWLIVFLTLCCG-----LGLLFA 518  
Db 622 RSLLOYGGSANRESV-----QGVTPH-----LAAQGHAEWVALLLS 659  
QY 519 FTVHPNLGIAVSFLALLYIFIVYFGRREGESNNMAWVLSRLARHIGVLELLK-- 577  
Db 660 KOANGNLGNKSGLTPLHLV-----AQBG-----HVPVADVLIKH 693  
QY 578 -----LMFVNPPPELPEQTKALPVRLP-----TDYNRLLSVG-----GETSLA 616  
Db 694 GNMVDATTRGYTPLHVAHYGNIKLVFLQHOADVNAKTKLGYSPHQAQOQHTDIV 753  
QY 617 EMI-----ATLSDACERBFGFLA--TRLFRVFKTEDTOGKKWKTKCCLPSFV 662  
Db 754 TLLKNGASPNVSSDGTPLATAKRLGIYSVTDVLKVV--TDET-----SFV 799  
QY 663 IFLFIIGIISITLAIARFVDPKHL-----TNAVLIISASVVG---LAFVLCNCTWQ 714  
Db 800 L-----VSDKHRMSPFETVDEIL--DVSEDEBELISFKAERR----- 835  
QY 715 VLDSLLNSQKRLHNAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFTQNTRLVVII 774  
Db 836 --DSRDVDEKELDVFKLDQVVESPAIPRIPE-----AMPETVIRSEQOASKEY 897  
QY 775 DGLDACEQDKVLQMLDTRVRLFS--KGFPIAIPASDP-----HIK----- 814  
Db 888 DEDSLPSSPATETSDNISPAVSPVHTGFLVSPWDARGSMGRSHRGLRVIPRTCA 947  
QY 815 -----ANQNLNSVLRDSNNGHDYMRN--IVHLPVFL--NBRGLS 851  
Db 948 APTRITRLVKPQKLSPTPPPLAEEGLASRIIALGPTGAQFLSPVIVEIPHAFSHGRG-- 1005  
QY 852 NARKFLVTSATNGDVPCSDTTGQEDADRVSQNSLGEWTKLGSKTALNRDTRYRQWQ 911  
Db 1006 -DRELVLVRSNGSV-----WKEHRSRYGESYLDQILN--GMDEELGSLSELEKRV 1055  
QY 912 RTITRQMSFDLTLLVLT---EDWFSDISPOTMRLLNIVSVTGRLLRANQISFNWDLAS 968  
Db 1056 RIIT--TDFPLYFVMSRLCOD--YDIGPEG-----GSLKSLVPLVQATPENAVTK 1105  
QY 969 WINLTQMPYRTSWLILYLEETEGIPDQWTLK-----TIYERISK----- 1008  
Db 1106 RVKLAQ-----AQVPDELVTKLGNQATPSPIVTPVPRRRKHPRIGLR 1151  
QY 1009 -NIPYTKOVPLEIDGDIRNFEV-----LSSRTPLVAVDVKVFLPCTV 1053  
Db 1152 IPLPWSMTNDRSGEDGTTSLRLCSVIGGTDOAQWEDITGTYKLYANECANF---TT 1208  
QY 1054 NLDPKLREIADYRAAREQISIGGLAYPPLPLHEGPPRAPSQSPSVCSSTSFNGPFA 1113  
Db 1209 NV--SARFWLSDCFRTAEAVNPATLYKEL-----TAPYMAKFV 1246  
QY 1114 GGVSQPHSSYSYSGMTGPHQPFYNRGSGPAGPV--VLLNSLNVDAVCE-----K 1162  
Db 1247 -----IPAKMNDPREGLRCYCMTDDKVDKLTLEQHENFVEVAR 1284  
QY 1163 LKQIEGLDQSWLPQYCTTIKKANINGRVLQAQCNIDELKEMNNVFGDWHLFRSTVLEMRN 1222

Db 1285 SRDIEVLECMSL-----FAELSG-----NLVPVKAAQORSFHFQSFRENRLAM-- 1328  
QY 1223 AESHVVPEDPRFLSSSSSGPAPHGEPAPASNE-----LPHTELS----- 1263  
Db 1329 -----PVKVRDSSREPFGGSLSFRLKAMKYEDTQHILCHLNITMPPCAKSGAEDRR 1379  
QY 1264 -SOTPYTNFNSFEELNTLGLDEGAPR-----HSNLSWOSQOTRRTP--SLSSLSNQD 1311  
Db 1380 RPTPLALRYSILSSTPGSLSGTEQAEMKMAVISEHLGLSWAELARELOFSVEDIN--- 1436  
QY 1312 SSIEISKLTDKVQAEYRDAYRIAMSQL-----EGGPGSTTISGRSSPHSTYIMQSSS 1367  
Db 1437 -----RIRVENPNSLLEQSVALLNLWIREG-----QVANMENLYTALQSID 1478  
QY 1368 GSGIHSNLFQEKQKOSBPDDGRKSFMLKMGVDIDYSSGVSSTNDASPLDPIDEDE-- 1425  
Db 1479 RGEIVNMLEGSRQSRNLKPD-----RRHTDRDYSLSPSQMGNGYSSL-----QDELL 1525  
QY 1426 -----KSDOSGSKL-----LPGKSSERSSLFQTDLKLKSGSLRYQKLPSDE 1467  
Db 1526 SPASLGALSSPLRADQYWNVAVLDAIPLAATEHDTMLESDMQVWSAGLTPSLVTAED 1585  
QY 1468 DE---SGTESONTPLKDDKDKAEKGVKVERVPKSPHSAEPIRTPIKAKEYLSDALLOK 1524  
Db 1586 SSLECSKAEDSAT-----GHEWKLEGALSERPGPE-----LGSILEVEDDVTDS 1631  
QY 1525 KDS-----DSGVRSSESPNHLNHEVADDQLEKANLIELEDDSHSGKRGPHSL 1576  
Db 1632 DATNGLIDLLEQEBEGORSEKLPGSKRQDDATGAGQ--DSENEVSLVSGHQRGQARITHS- 1689  
QY 1577 SGLQDPIIARMSICSDEK-----KSPSECSLIASSPEENW 1611  
Db 1690 -----PTVSQVTERSQDRLQDWDADGSIVSYLQDAAQGSW 1724

## RESULT 2

US-09-949-016-9011  
; Sequence 9011, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9011  
; LENGTH: 1883  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9011

Query Match 6.3%; Score 557.5; DB 4; Length 1883;  
Best Local Similarity 19.7%; Pred. No. 1.9e-38;  
Matches 374; Conservative 286; Mismatches 659; Indels 581; Gaps 72;

QY 21 LKALEKCKDVBNERCGOTPLMIAAEOGNLEIVKELIKNGANCNLEDNMTALISASK 80  
Db 97 VRELNVYGANVAQSQKGTPLYMAAQENHLEVVKELLENGANQNVATEDGFTPLVALQ 156  
QY 81 EGHVHIVEELLKCGV-----NLEHRDMGQWATLMWA 111  
Db 157 QGHENVVAHLINVTGKVRKLPALHIAARNDDTRTAAVLLQNDPNPDVLSKTGFTPLHIA 216  
QY 112 CYKGRDVTVELLSHGANSVT-----GLQY-----SV 139



Db 217 AHYENLANVAQLLNRGASVNFPTQNGITPLHIAARRGNVIMVRLLDGCAIETKTDEL 276  
Qy 140 YPIIWAAGRHADIVHLLQNGAKVNSDKYGTTPLYWAARKGHLEKCHLLAGADVD- 198  
Db 277 TPLCHAARNGHVRISIELLDHGAPQIAKTRNGLSPIHWAAGDHLQVLLQYDAEIDD 336  
Qy 199 -----OECA-----NSMTALIVAVKGYTQSVKEILKR 226  
Db 337 ITLDHLTPHVAACHGHRHVAKVLDDKGAKNPNSALNGFTPLHIAACKXNVVRVWELLJKT 396  
Qy 227 NPNVNLTKDQNTALMIASKEGHTETIYODLLDAGTYVNIPIRSGDTVLIGAVRGHVEIV 286  
Db 397 GASIDAVTESGLTPLHVASFNGHLPVIVKNLLQRGASPNVSVKVVETPLHMAARAGHTEVA 456  
Qy 287 RALLQKQVADIDIRGDNKTALYMAVEKGNATVMDILQCNPDTEICTKGETPLIKATKM 346  
Db 457 KYLLQNKAVNAKADODTPLHCAARIGHTVNVKLLNANPNLATTAGHTPLHIAARE 516  
Qy 347 RNIWEVLLLDKGAKSAVDKKGOTPLHIAIRGRSRKLAELLRNPKDGRLLYRPNKAGE 406  
Db 517 GHVETVALLLEKEASQACMTKKGFTPLHVAKYKVRVAELLLR- RDA-----HPNAAKG 570  
Qy 407 ---TPYNDCHQKSIITQIFGARHLSPTETDGMILGY-DIYSSALADIILEPTMQPPIC 462  
Db 571 NGLTPLHVAVHHNNLDIVKLLPRGGSPHSPAWN--GYTPLHIAAKQN-----QVEVA 621  
Qy 463 VGLYAQWGSKGSFLKKLEDEMTKFPAGQIPLFPQSWLIIVFLTLCCGG-----LGLLFA 518  
Db 622 RSLQYGGSANAEV-----QGVTPH-----LAAQEGHAEVVALLS 659  
Qy 519 FTVHPNIGIAVSLFLALLYIFIVIPGGRREGESWNAWVLSTRLARHIGYLELLK- 577  
Db 660 KQANGNLGNKSGLTPLHLV-----AQEG-----HVPVADVLIXH 693  
Qy 578 -----LMFVNPELPEQTKALPVRFLP--TDYNRJLSSVG-----GETSLA 616  
Db 694 GVMVDATTMGYTPHVAHYGNIKLVKFLIHOQADVNAKTKLGYSPLHQAQOQHDTIV 753  
Qy 617 EMI-----ATLSDACEREGFLA-TLFRVFKTEDTQGGKWKTKCLSFV 662  
Db 754 TLLKNGASNEVSSDGTPLAIARLIGYSIVTDVLKV-TDET-----SFV 799  
Qy 663 IFLEFIIGIISGITLLAIFRVDPKHL-----TVNAVLIASVVG-----LAFVLCRTWQ 714  
Db 800 L-----VSDGRMSFPETVDEIL-DYSEDEGEELISFKAER-----835  
Qy 715 VLDSLLNSQRKLHNAASKHLKSEGMKVLKCEVELMARMAKTIDSTFQNTQRLVVI 774  
Db 836 --DSRDVDEEKELLDVFKLDQVVESPAIPRIPC-----AMPETVIVIRSEEQEQAKEY 887  
Qy 775 DGLDACQDKVLQMLDTRVRLFS--KGFPTAIFASDP-----HIILK-----814  
Db 888 DEDSLIPSPATETSDNISVASPVHTGFLVSMVDARGSMGRSNGRLVRPIPRPCA 947  
Qy 815 -----AINQNLNSVLDRSNDINGHYMRN-IVHLPVFL-NSRGLS 851  
Db 948 APTFITCLVKPKLSTPPPLAEGLASRIIALGPTGAQFLSPVIVEIPHFAHSGRG-- 1005  
Qy 852 NARKFLVTSATNGDVPSCDTHGIEDADRVSQNSLGEMTKLGSKTALNRDRDTRRRQMQ 911  
Db 1006 -DRELVLVRSENGSV-----WKEHRSRYGESYLDQILN-GMDEBELGSLELEKRCVC 1055  
Qy 912 RTITROMSFDTLLLVLT---EDWFSDISPQTMRLNIVSVTGELLRANQISFNWDLAS 968  
Db 1056 RIIT--TDFPLYFIMRLQOD-YDTIGPEG-----GSLKSLVPLVQATPENAVTK 1105  
Qy 969 WINLTEQWPYKTSWLIYLEETEGIPQWTLK-----TIVERISK-----1008  
Db 1106 RVKLALQ-----AQVPDELVTLLGNQATFSPITVVEPRRKHFRPIGLR 1151  
Qy 1009 -NIPTTKDVPFLBIDGDIRNFVF-----LSSRTPVLVARDVKVFLPCTV 1053

Db 1152 IPLPPSTNDNRDSGEQDTISRLLLCSVIGTDOAQWEDITGTTKLIVYANECANF---TT 1208  
Qy 1054 NLDPKLEIADYVAAAREQISIGGLAYPPLPLHGGPPRAGSGYQPSVCSSTSFNGPFA 1113  
Db 1209 NV--SARFWLSDCPRTAENVFATLLYKEL-----TAVPYMAKPV 1246  
Qy 1114 GGVVSPHSHSYSGMTGPQHPFYNRGSGPAGPV--VLLNSLVNDVACE-----K 1162  
Db 1247 -----IPAKVNDPREGRLRCVCMWDDKVDKLEQHENFVEVAR 1284  
Qy 1163 LKQIEGLDQSMPLPOYCTTIKANKINGRVLACQNDIDELKKNMFMFGMHLPRSTVLEMRN 1222  
Db 1285 SRDIEVLEHMSL-----FAELSG-----NLVPVKAAQQRSFHQSFRENRLAM-- 1328  
Qy 1223 AESHVVPEDPRFLSESSGPAPHCEPARRASHNE-----LPHTELS-----1263  
Db 1329 -----PKVRDSSREPGGSLSFURKAWKYEDTOHILCHLNIWTPPCAKGSGAEDRR 1379  
Qy 1264 -SOTPYTINFSEELNTLGLDEGAPR-----HSNLSWQSQTRTP--SLSSLNSQD 1311  
Db 1380 RTPPLALRYSILSESTPGSLSGTEQAEMKMAVISEHLGLSWAELARELQFVEDIN--- 1436  
Qy 1312 SSIEISKLTQVQAYRDAYREYIAQMSQL-----EGPGSTTISGRSPHSHYMGQSSS 1367  
Db 1437 -----RIRVENPNSLLEQSVALLNLWVIREG-----QANMENLYTALQSID 1478  
Qy 1368 GGSITHSLNEQEKQKSEPKDDGKRSFLMKRGVDVIDYSSGVSTNDASPLDPITEEDE-- 1425  
Db 1479 RGEIVNMLEGSGQSRNLKPD-----RRHTRDYSLSPSQWNGYSSL-----QDELL 1525  
Qy 1426 -----KSDQSGSKL-----LPQKKSERSSLFOTDLKLGSGLYQKLPSEDE 1467  
Db 1526 SPASLGCALSSPLRADQYVNEVAVLDAIPLAATEHDTMLEMSDMQVWSAGLTPSLVTAED 1585  
Qy 1468 DE---SSTESDNTPLKDKKAKGKVRVPKSPHSAEPIKFTFAKEYLSDALLDK 1524  
Db 1586 SSLECSKAEDSAT-----GHEWKLEGALSEEPGRPE-----LGSLELVEDDTVDS 1631  
Qy 1525 KDSS-----DSGVRSSESSPNHSLHNEVADDSOLEKANLIELEDDSHSGKGIHSL 1576  
Db 1632 DATNGLIDLEQEGQKSEKLPQSKRQDDATGAGQ--DSENEVSLVSHGORGQARITHS- 1689  
Qy 1577 SGLQDPFIARMSICSEDK-----KSPSECSLIASSPEENW 1611  
Db 1690 -----PTVSQVTERSQDLQDWDADGSIYVLQDAAGSW 1724

RESULT 3

US-09-949-016-9012  
; Sequence 9012, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9012  
; LENGTH: 1883  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9012  
Query Match 6.3%; Score 557.5; DB 4; Length 1883;





; Sequence 9014, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9014  
; LENGTH: 1883  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-9014

Query Match 6.3%; Score 557.5; DB 4; Length 1883;  
Best Local Similarity 19.7%; Pred. No. 1.9e-38;  
Matches 374; Conservative 286; Mismatches 659; Indels 581; Gaps 72;

QY 21 LKALLEKCDVDERNCGOTPLMIAAQQNLEIVKELIKNGANCNLEDLDNWTALISASK 80  
DB 97 VRELNVNGANVAQSQKGFPLYMAAQENHLEVVKELLENGANQNVATEDGFTPLAVALQ 156  
QY 81 EGHVHIVBELLKGV-----NLEHRDMGWTALMWA 111  
DB 157 QGHENVVAHLINTGKGVKRLPALHIAARNDTRTAAVLLQNDPNPDVLSKTGFTPLHIA 216  
QY 112 CYKGRDVBELLSHGANPSVT-----GLQY-----SV 139  
DB 217 AHYENLVAQLLNRGASVNFQNGITPLHIAARRGNVIMVRLLDGAGQIETKXDEL 276  
QY 140 YPIIWAAGRHADIVHLLQNGAKVNCSDKGTTPVLVMAARKHLECVKHLAMGADVD- 198  
DB 277 TPLHCAARNGHVRISLEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVRLLLQYDAEIDD 336  
QY 199 -----QEGA-----NSMTALIVAKGGYTSQVKELKR 226  
DB 337 ITLDHLTPLHVAACHGHRVAKVLLDKGAKPNRSLNGFTPLHIAACKONHVRVWELLKT 396  
QY 227 NPNVNLTDKDGNTALMASKEGHEIIVQDLIDAGTYVNIPIRSGDVTVLIGAVRGHVEIV 286  
DB 397 GASIDAVTESGLTPLHVASFWGHLPIVKNLLQRGASPNVSNVKTPLHMAARAGTEVA 456  
QY 287 RALLQKYADIDIRQDNKTALYWAVEKGNATWVRDILQCNPDTEI CTKQGETPLIKATKM 346  
DB 457 KYLLQNKAKVNAKADQDTPLHCAARIGHTNMVKLLLENNANPLNATATTAGHTPLHIAARE 516  
QY 347 RNTEVVELLDKGAKSADVKKGDTPLHIAIRGSRKLAELLILRNPKDGRLLYRPNKAGE 406  
DB 517 GHVETVLALLEKASQACMTKGTGFTPLHVAAKYGVKRVAEELLE--RDA----HPNAAGK 570  
QY 407 ---TPVNICDSHQKSILTOIFGARHLSPTETDGMGLGY-DLYSALADILSEPTMOPPIC 462  
DB 571 NGITPLHVAHNNLDIVKLLPRGSGSPHSPAWN--GYTPLHIAAKQN-----QVEVA 621  
QY 463 VGLYAQWGSKSFLLKKLEDEMKTFAQQIIEPIFQFSWLVIFVTLILCCG-----LGLLFA 518  
DB 622 RSLQYGGSANAEV-----QGVTPH-----LAAQEGHAEMVALLIS 659  
QY 519 FTVHPNLGIAVLSFLALLYIFPIVYFGRRGESNNAWVLSLRLARHIGVLELLK- 577  
DB 660 KOANGNLGNKSGLTPLHLV-----AQEG-----HVPVADVLIKH 693  
QY 578 -----LMFVNPPELPEQTTRKALPVRFLLF-----TDYNRLSSVG-----GETSLA 616

694 GVMVDATRMGYTPLHVASHYGNIKLVKELLQHOADVNAKTKLGYSPHQAQQOCHTDIV 753  
617 EMI-----ATLSDACERBEFGFLA-TRLPFRVFKTEDTQKKKKWKTCCCLPSFV 662  
754 TLLKNGASPNVSSDGTTPLAIAKRLGYISVTDVLKV-TDET-----SFV 799  
663 IFLFIIGCIISGITLLAIFRVPDKHL-----TNAVLISIASVVG--LAFVLNCRWTWQ 714  
800 L-----VSDKHRMSPETVDEIL-DVSEDEGELSFFKAERR----- 835  
715 VLDSLINSQRKRLHNAASKHLKSEGFMKVLKCEVELMARMAKTIDSTQNTQRLVIL 774  
836 --DSRDVDEKELLDFVPLQDVVESPAIPRIPC-----AMPETVIRSEEQEQASKEY 887  
775 DGLDACEQKVQLMDTVRVLFS--KGPIAIFASDP-----HIILK----- 814  
888 DEDSLIPSSPATETSDNISPVASPVHTGFLVFMVDARGSMRSGRHLNGLRVVIPPRTCA 947  
815 -----AINQNLNSVLKRDNSINGHDYMRN-IVHLPVFL-NSRGLS 851  
948 APTRITCRVLKPKOKLSTPPPLAEEGLASRIIATGFTGAQFLSPVIVEIPHFAHSGRG-- 1005  
852 NARKFLVTSATNGDVPSCDITGIEDADRRVSNQSLGENTKLGSKTALNRRDITYRRQMQ 911  
1006 -DRELVLRSNGSV-----WKEHRSRYGESYLDQILN-GMDEELGSLSELEKRCVC 1055  
912 RIITROMSFDLTKLVT---EDWFSDISPOTWRRLLNIVSVTGRLLRANQISFNWDLAS 968  
1056 RIIT--TDFPLFYVIMSRQCQD-YDTIGPEG-----GSLSKSLVPLVQATPENAVTK 1105  
969 WINLTSQMPYRTSWLLILYEETEGIPDOMTLK-----TIVERISK----- 1008  
1106 RVKLAQ-----AQVPELVTLLKLGNOATFSPVIVPEPRRKHRRIGLR 1151  
1009 -NIPTTKVPELLEIDGDIRNFEVF-----LSRTPVLVARDVKVFLPCTV 1053  
1152 IPLPSSWTDNPRDSGEGDITSLRLCLSVIGGTDQAWEDITGTTKLIVYANECANF--TT 1208  
1054 NLDPKLEIADVRAAREQISIGGLAYPLPLHEGPPRAPSGVSPQSPVCSSTSFNGPFA 1113  
1209 NV--SARFWSLDCPRTAEAVNFATLLYKEL-----TAVPYMAKFV 1246  
1114 GGVSVPQHSSVYVSGMTGPQHPFYNRSGSPAGPV--VLLNSLVNDVAVCE-----K 1162  
1247 -----IFAKNDPREGLRCYCMTDKDKVKTLEOHENFVEVAR 1284  
1163 LKQIEGLDQSMPLPQYCTTIKKANINGRVLAQCNIDELKEMNNMFGDWHLFRSTVLEMRN 1222  
1285 SRDIEVLEGMSL-----FAELSG-----NLVPVKAAQQRSPHFQSFRENRLAM-- 1328  
1223 AESHVVPDPRFLSESSSGPAPGEPARRASHNE-----LPHTELS----- 1263  
1329 -----PVKVRDSREPGSLSFLRKAMKYEDTQHILCHLNIITMPPCAKSGAEDRR 1379  
1264 -SOTPYTLNFSPEELNLTGLDEGAPR-----HSLNSQSQTRTP-SLSSNSQD 1311  
1380 RPTPLALYSLISESTPGSLSGTEQAEMKMAVISEHLGLSWALARELQFVEDIN--- 1436  
1312 SSIEISKLTDKVQAEYRDAYREYIAQMSQI-----EGPGSTTISGRSSPHSTTYMGQSSS 1367  
1437 -----RIRVENPNSLLEQSVALLNLVIREG-----QANMENLYTALQSID 1478  
1368 GGSIHNSLHQEKGDSEPKPDGGRKSFMLKRGDVIDYSSGVSTNDASPLDPIEDE-- 1425  
1479 RGEIVNMLEGSGRQSNLKP-----RRHTDRDYSLSPSQMGVYSSL-----QDELL 1525  
1426 -----KSPQSGSKL-----LPKKSSESSSLFQTDLKLKGLRKYQLKPSDE 1467  
1526 SPASLCALSSPLRADQYNEVAVLDAIPLAATEHTMTLEMSDMQVWAGLTSVLTAE 1585  
1468 DE---SGTBESONTPLLKDDKQKAEKGVKVERVPKSPHSAEPIRTTFIKAEYLSALLDK 1524  
1586 SSLECSKAEDSDAT-----GHEWKLSEALSEEPRGPE-----LGSLELVEDDVTDS 1631



QY 1368 GGSIHNLBOEKQKSEPKDDGKRSFLMKRGVDIYSSGCVSTNDASPLDPTTBEDE-- 1425  
DB 1479 RGEIVNLEGSGRNLKPD-----RRHTRDYSLSPSQMNGYSSL-----QDELL 1525  
QY 1426 -----KSDQSGSKL-----LPCKKSSRSSLFQTDCLKGSGRLRYQKLPSDE 1467  
DB 1526 SPASLGALSSPRAOYVNEVAVLDAIPLAATEHTMLEMSQVWSAGLTSLVTAED 1585  
QY 1468 DE---SGTESDNTPLKDKKORAKGKVERVPKSPHSAEPIRTTIFIKAKYLSDALDK 1524  
DB 1586 SSLECSKAEDSDAT-----GHEKLEGALSEEPRGPE-----LGSLELVEDDVTDS 1631  
QY 1525 KDS-----DSGVRSSSPHSLNEVADDSQLEKANLIELEDDSHSGKRGPHSL 1576  
DB 1632 DATNGLIDLLEQEGORSEKLPGRQDDATGAGQ--DSNEVSLVSGHQRGOARITHS- 1689  
QY 1577 SGLQDPFIARMSICSEDK-----KSPSECSLIASSPEENW 1611  
DB 1690 -----FTVSQVTERSQRLQDWDADGSIYSYLQDAAGSW 1724

RESULT 7  
US-09-949-016-9016  
; Sequence 9016, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9016  
; LENGTH: 1883  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9016

Query Match 6.3%; Score 557.5; DB 4; Length 1883;  
Best Local Similarity 19.7%; Pred. No. 1.9e-38;  
Matches 374; Conservative 286; Mismatches 659; Indels 581; Gaps 72;

QY 21 LKALLECKDVERNECGQTPMLAAEQGNLEIVKELIKNGANCNLEDLNDWTALISASK 80  
DB 97 VRELNVYGANVAQSQKGFTPLYMAAQENHLEVVKFLENGANQNVATEDGFTPLAVALQ 156  
QY 81 EGHVHIVEELLKGV-----NLEHRDMGGWTALMWA 111  
DB 157 QGHENVVAHLINYGTKGVRLPALHIAARNDTRTAVALQNDPNPDVLSKTGFTPLHIA 216  
QY 112 CYKGRDVBVELLSHGANSVPT-----GLQY-----SV 139  
DB 217 AHYENLVAQLLNRRGASVNTFQNGITPLHIASRRGNVIMVRLLDRGAIQETKDEL 276  
QY 140 YPIIWAAGRHADIVHLLQNGAKVNGSKYGTGTPPLVMAARKGLECVKHLAMGADVD- 198  
DB 277 TPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVRLLLQYDAEIDD 336  
QY 199 -----QEGA-----NSMTALIVAKGGYTSQVKEILKR 226  
DB 337 ITLDHLTPLHVAACHGHRVAKVLLDKGAKPNRSLANGFTPLHIAACKNHRVWELLKKT 396  
QY 227 NPNVNLTKDGNLALMASKEGHEIIVQDILLDAGTYVNIPIDRSDGTVLIGAVRGHVEIV 286

DB 397 GASIDAVTESGLTPLHVASFMGHLPIVKLLQRGASPNVSNVKVETPLHMAARAGHTEVA 456  
QY 287 RALLQKYADIDIRGQNKNTALYVWEKGNATWVRDILQCNPDTEICTKOGTEPLIKATKM 346  
DB 457 KYLLQWKAQVNAKADQDTPHCAARIGHTNMVKLLLENNANPLATTAHTPLHIAARE 516  
QY 347 RNIEVELLIDGAKVASVDKGDTPHIAIGRSRKLAEILLRNPKDGRLLYRPNKAGE 406  
DB 517 GHVETVLALEKEASQACMTKKGFTPLHVAARYKVRVAELLE--RDA-----HPNAAKG 570  
QY 407 ---TPYINIDCSHQKSLTQIFGARHLSPTETDGDMLGY-DLYSSALADILSEPTMOPPIC 462  
DB 571 NGITPLHVAHNNLDIVKLLPRGSGSPHAWN--GYTPLHIAAQN-----QVEVA 621  
QY 463 VGLYAWGSGKGFLLKKLEDEMKTFAQOIEPLFQFSWLIVFLTLCCG--LGLLFA 518  
DB 622 RSLQYGGSANAEV-----QGVTPH-----LAAQEGHAEMVALLS 659  
QY 519 FTVHPNIGIAVSLFLALLYIPFVIYFGRRGESNNAWVLSRLARHIGVLELLK- 577  
DB 660 KQANGNKGNSGLTPLHLV-----AQEG-----HVPVADVLK 693  
QY 578 -----LMFVNPPPELPEQTTKALPVELF---TDYNRLSSVG-----GETSLA 616  
DB 694 GUMVDATTRMGYTPHVAHYGNIKLVKELLQHQADVNAKTGLGYSPLHQAAQOQHTDIV 753  
QY 617 EMI-----ATLSDACEREFGLA--TRLFRVFKTEDTQKKKKWTKCCLPSFV 662  
DB 754 TLLKNGASPNVSSDGTTPLAIAKRLGYISVTDVLKV--TDET-----SFV 799  
QY 663 IFLFIGIISGITLLAIFRVDPKL-----TVNAVLSIASVVG--LAFVLNCRTWQ 714  
DB 800 L-----VSDKHRMSPETVDEIL-DVSEDEGEELSFKARR-----835  
QY 715 VLDLNLNQRKRLHNAASKLHLKSEGFMKVKCEVELMARMAKTIDSTQNTQRLVVI 774  
DB 836 --DSRDVDEKELDFVKLDQVVESPAIPIC-----AMPETVIRSEEEQASKEY 887  
QY 775 DGLDACEQDKVQMLDVTIVLFS--KGFPIAFASDP-----HIILK-----814  
DB 888 DEDSLIPSSPATETSDNISPVASPVHTGFLVSPMVDARGSMRSGRHLNGLRVVPIPTCA 947  
QY 815 -----AINONLNSLRDSDNNGHDYMN-IVHLPVFL-NSRGLS 851  
DB 948 APTRITCRLVKOKLSTPPPLAEEGLASRIALGPTGAQFLSPVIVEIHPHSHGEG-- 1005  
QY 852 NARKFLVTSATNGDVPSCDSTTIGIEDADRRVQNSLGMTKLGSKTALNRRDITYRRQMQ 911  
DB 1006 -DRELVLVRSNGSV-----WKEHRSRYGSEYLDQILN-GWDEELGSLSELEKKRVC 1055  
QY 912 RTITROMSFDLTKLVLT---EDWFSDISPOTMRRLLNIVSVTCRLLRANQISFNWDELAS 968  
DB 1056 RIIT--TDFPLXFMVSRLLCQD--YDTIGPEG-----GSLKSLVPLVQATFPENAVTK 1105  
QY 969 WINLTQMPVRTSWLLIYLEETEGIPDOMTLK-----TIVERISK-----1008  
DB 1106 RVKLAQ-----AQPVDELVTKLLGNOATFSPITVPEPRKKRHPRIGLR 1151  
QY 1009 -NIPTTKQVEPLEIDGDIRNFEVF-----LSSRTPVLVARDVKVFLPCTV 1053  
DB 1152 IPLPSPWNTNPRDSGEGDITSLRLCSVIGTDAQWEDITGTTKLIVYANECANF---TT 1208  
QY 1054 NLDPKUREIADVRAAREQISIGGLAYPPLPLHEGPPRAPSGYQSPVSCSSYFNGPFA 1113  
DB 1209 NV--SARFMLSDCPRTAEAVNFATLLYKEL-----TAVPYMAKFV 1246  
QY 1114 GGVSYPQHSSYYSGMTGPHQHPYNRGSGPAGPV--VLILNSLVDAVCB-----K 1162  
DB 1247 -----IFAKNDPREGRLCYCMTDDKDKVTKLEQHENFVEVAR 1284  
QY 1163 LKQIEGLDOSMLPOYCTTIKANINGRVLAAQCNIIDELKEMNNMFGDHLFRSTVLEMRN 1222  
DB 1285 SRDIEVLGMSL-----FAELSG-----NLVPVTKAAQQRSHFQSFRENRLAM-- 1328



Qy	1223	AESHVPEPDRFLSGSSGAPHGEPARRASHNE-----LPHTELS-----	1263
Db	1329	-----PVKVRDSSREPGLSPLRKANKYEDTQHILCHLNIWTPPCAKSGAEDRR	1379
Qy	1264	-SQPYTLNPSFELNITGLDEGAPR-----HSNLSWQSQTRTP-SLSLNSQD	1311
Db	1380	RTPTFLALRYLSILSESTPGSLSGTEQAEKMAVISEHLGLSLWABELARELOPVEDIN---	1436
Qy	1312	SSIBISKLTDKVAEYRDVREYIAQMSQL-----EGGPGSTTISGRSPHSSTYMGSSS	1367
Db	1437	-----KIRVENPNSLLEQSVALLNLWVIREG-----QNMENMLYALQSID	1478
Qy	1368	GGSHNSLEQKGDSPKPDGGRKSLFKRGDVIDYSSSGSVNDASPLDPITEEDE--	1425
Db	1479	RGEIVNLEGGSGROSLKPD-----RHTDRDYSLSPQNGYSSL-----QDELL	1525
Qy	1426	-----KSDQSGSKL-----LPGKKSERSLSLFTDULKLGSLGRYOKLPDSDE	1467
Db	1526	SPASLGICALSPLRADQYWEVAVLDAIPLAATBHDITMLESDMQVWSAGLTPSLVTAED	1585
Qy	1468	DE---SGTESDNTPLKDDKORKAEGKVERVPKSPHSABEPIRTFIKAKEYLSDALDK	1524
Db	1586	SSLECSAEDSDAT-----GHEWKLEGALSEEPGRPE-----LGSLELVEDDTVDS	1631
Qy	1525	KDSS-----DSGVRSSSESSPNHSLHNEVADDSQLEKANLIBLEDSDSHGKRGIPHS	1576
Db	1632	DATNGLDILLEQEGQSGSEKLPQSKRQDDATGAGQ--DSENEVSLVSGHQRGQARITHS-	1689
Qy	1577	SGLODPIIARMSICSEDK-----KSPSECSLIASSPEENW	1611
Db	1690	-----PTVSQVTERSQRLQDWDADGSIVSYLQDAAQGSW	1724

RESULT 8

US-09-949-016-9017

Sequence 9017, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9017

LENGTH: 1883

TYPE: PRT

ORGANISM: Human

US-09-949-016-9017

Query Match

Best Local Similarity 19.7%; Pred. No. 1.9e-38;

Matches 374; Conservative 286; Mismatches 659; Indels 581; Gaps 72;

Qy	21	LKALLECKDVERNECGQPTLMAAEOGNLEIVKELIKGANCNLEDNLWTALISASK	80
Db	97	VRELNVGANVAQSQSGFTPLYMAAQENHLEVVKFLLENGANCNATEDGFTPLVALQ	156
Qy	81	EGHVHIVEELLKCV-----NLEHRDGMGWALMWA	111
Db	157	QGHENVVAHLINYGTKVKRLPALHIAARNDDRTAAVLLQNDPNPVLKGTGFTPLHIA	216
Qy	112	CKGRDVTWELLSHGANPSVT-----GLQY-----SV	139

QY 1054 NLDPKLEIIADYRAAREQISIGGLAYPPLPLHEGPRAPSGYSPVCSSTSFNGPFA 1113  
DB 1209 NV--SARFWLSDCPTAEAVNFATLLYKEL-----TAVPYMAKFV 1246  
QY 1114 GGVSPQPHSSYYSGMTGPHFPYNGSGPAPGV--VLINSLNVDACE-----K 1162  
DB 1247 -----IPAKNDPREGRLACYCMTDDKVKLTLEOHENFVEVAR 1284  
QY 1163 LKQIEGLDQMLPOYCTTIKANINGRVLACQNIDELKEMNNFGDWHLFRSTVLEMRN 1222  
DB 1285 SRDIEVLEGMSL-----FABELSG-----NLAVPKAAQORSFHFQSFRENRLAM-- 1328  
QY 1223 AESHVVPEDPRFLESSESSGAPGEGEPARRASHNE-----LPHTELS----- 1263  
DB 1329 -----PVKVRDSREPGGSLSPKSGTEQAEMKMAVISEHLGLSWAELARELOFSVEDIN--- 1379  
QY 1264 -SOTPYTLNFSPEELNLTGLDECAPR-----HSLNSWOSQTRTP-SLSSLSNSQD 1311  
DB 1380 RPTPLALRYSLISESTPGSLSGTEQAEMKMAVISEHLGLSWAELARELOFSVEDIN--- 1436  
QY 1312 SSIEISKLTKQVARYDAYREYIAQMSQI-----EGPGSTTISGRSSPHSTYYMGQSSS 1367  
DB 1437 -----RIRVENPNSLLEQSVALLNLWVIREG-----QANMENLYTALQSID 1478  
QY 1368 GGSIHNSLQEKQKDEPDDGKRSFLMKRGVDIYSSGVSSTNDASPLDPITEDE-- 1425  
DB 1479 RGEIVNMLEGSGQSRNLKPD-----RRHTDRYSLSPSQMGYSYL-----QDELL 1525  
QY 1426 -----KSPQSGSKL-----LPCKKSSRSLSFQTDCLKGLGRLYQKLPSPDE 1467  
DB 1526 SPASLGCALSSPLRADOYNEVAVLDAIPLAATEHTMTLEMSQWVSAGLTPSLVTAED 1585  
QY 1468 DE---SGTESDNTPLKDDKDKARKGKVERVPKSPHSAEPIRTPIKAKEYLSDALDK 1524  
DB 1586 SSLEKSAEDSDAT-----CHEWKELGALSEEPRGP-----LGSLELVEDDTVDS 1631  
QY 1525 KOSS-----DSGVRSESSPHSLNEVADDSQLEKANLIEDEDDSHSGRGKIPHSL 1576  
DB 1632 DATNGLIDLLEQEGORSEKLPKSGKRDQDAGQ--DSENEVSLVSGHQGQARITHS- 1689  
QY 1577 SGLQDPPIIARMSICSEDK-----KSPSECSLIASSPEENW 1611  
DB 1690 -----PTVSQVTERSQDRQDWDADGSIYSYLQDAAQGSW 1724

## RESULT 9

US-09-949-016-6964  
; Sequence 6964, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6964  
; LENGTH: 1856  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6964

Query Match 6.3%; Score 555.5; DB 4; Length 1856;  
Best Local Similarity 19.7%; Pred. No. 2.7e-38;

Matches 374; Conservative 286; Mismatches 659; Indels 581; Gaps 72;  
QY 21 LKALLEKCKVDNERNECGOTPLMIAAEOGNLEIVKELIKNGANCNLEDLNDWTALLSASK 80  
DB 94 VRELNYGANVAQSQGFTPLYMAAQENHLEVYKFLLENGANQNATEDGFTPLVALQ 153  
QY 81 EGHVHIVBELLLKGV-----NLEHRDMGWTALMWA 111  
DB 154 QGHENVVAAHLINVTGKVRLLPALHIAARNDTRTAARVLLQNDPNDVLVSKGTFTPLHIA 213  
QY 112 CYKGRDVTVELLSHGANSVT-----GLQY-----SV 139  
DB 214 AHVENLVAQLLNRGASVNFQNGITPLHIASRRGNVIMVRLLLDRGQIETKTDEL 273  
QY 140 YPIIWAAGHADIIVHLLONGAKVNCSDKYGTTPLVWAARKHLESCVXHLLANGADVD- 198  
DB 274 TPLHCAARNGHVRISILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVLRLLLOYDAEIDD 333  
QY 199 -----QEGA-----NSMTALIVAKGGYVTSQVKEILKR 226  
DB 334 ITLDHLTPLHVAHCHHRVAKVLLDKGAPNSRALNGFTPLHIACKKONHVRVMEILLKT 393  
QY 227 NPNVNLTKDGNATMIASKEGHEITVQDLDAGTYVNIIPDRSGDTVLICAVRGHVEIV 286  
DB 394 GASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKTPLHMAARAGHTVA 453  
QY 287 RALLQYADIDIRQDNKTALYWAKEGNATWVRDILQCNPDTEICTOGTEPLIKATKM 346  
DB 454 KYLLQKAKVNAKADQDQTPHCAARIGHTNMVKLLLENNANPNLATATTAGHTPLHIAARE 513  
QY 347 RNEVVELLLDKGAKVAVDKGDTPLHTAIGRSRKLAEILLRNPKDGRLLVRPNKAGE 406  
DB 514 GHVETVALLEKEASQACMKKGFTPLHVAKYKVRVAELLE--RDA-----HPNAAKG 567  
QY 407 ---TPYINDCSHQKSLTQIFGARHLSPTETDGMGLY-DLYSSALADILSEPTMOPPIC 462  
DB 568 NGITPLHVAHVNLDIVKLLPRGSGSPHSAWN--GYTPLHIAKQN-----QVEVA 618  
QY 463 VGLIYAQWGSGKFLKLEDEMKTFAQQIPLFPQSWLIVFTLLLCGG-----LGLLFA 518  
DB 619 RSLQYGGSANAEV-----QGVTPH-----LAAQEGHAEMVALLS 656  
QY 519 FTVHPNIGIAVSLFLALLVIFPIVIYFGRREGESMNWAVLSTRLAHIGVLELLK- 577  
DB 657 KOANGNLGNKSLTPLHLV-----AQEG-----HVPVADVLIK 690  
QY 578 -----LMFVNPPPELPRQTTKALPVRELF--TDYNRLSSVVG-----GETSLA 616  
DB 691 GVMVDATTRMGYTPLHVAHYGNIKLVKELLQHQADVNAKTGLGYSPHQAQOQHTDIV 750  
QY 617 EMI-----ATLSDACEREFGLA--TRLFRVFKTEDTQKKKWKTCCLPSPFV 662  
DB 751 TLLKNGASPNVSSDGTTPLAIAKRLGYISVTDVLKV--TDET-----SFV 796  
QY 663 IFLFIIGCIISGITLLAIFRVPKHL-----TVNAVLSIASVVG--LAFVLNCRWTWQ 714  
DB 797 L-----VSDKHRMSPFETVDEIL-DVSEDEGEELISFKAERR----- 832  
QY 715 VLDSLINSQKRLHNAASKHLKSEGFMKVLKCEVELMARMAKTIDSTQNTQRLVIVII 774  
DB 833 --DSRDVDEKELLDFVPKLDQVVESPAIPRIPC-----AMPETVIRSEEQEQASKEY 884  
QY 775 DGLDACEQKVIQMLDVTVRVLS--KGPFIAPASDP-----HIILK----- 814  
DB 885 DEDSLIPSSPATETSDNISFVASPVHTGFLVFMVDARGSMRSGRSHNGLRVVIPPRTCA 944  
QY 815 -----AINONLSNRDSDNINGHDYMRN-IVHLVPVFL-NSRGLS 851  
DB 945 APTITCRVLVKQKSLTPPPLEAEEGLASRIIATGTLGAQFLSPVIVEIPHFAHSGRG-- 1002  
QY 852 NARKPLVTSATNGDVPDSDTTGIEDADRRVRQNSLGMTKLGSKTALNRDRYRRROMQ 911  
DB 1003 -DRELVLRSNGSV-----WKEHRSRYGESYLDQILN-GMDEELGSLSELEKRVK 1052

Qy 912 RTITRQMSFDITKLLVT---EDWFSDISPOTMRLLNIVSVTGRLLRANQISFNWDRLAS 968  
Db 1053 RIIT--TDFPLYFVIMSLCQD-YDIIGPEG-----GSLKSLVPLVQATFPENAVTK 1102  
Qy 969 WINITEQWPYTSWLLIYLETEGIPDQWTLK-----TIVERISK----- 1008  
Db 1103 RVKLALQ-----AQPVDSLVTKLLGNQATFPSPVITVEPRRKRFRPIGLR 1148  
Qy 1009 -NIPTTKDVEPLLEIDGIRNFEVF-----LSRTPVLVARDVKVFLPCTV 1053  
Db 1149 IPLPSTNDPRDSEGGDTTSRLILCSVIGTDOAQWEDIYGTTLVYVANCANP---TT 1205  
Qy 1054 NLDPKLREIITADVRAAREQISIGGLAYPPLFHEGPPRAPSGYSGQPSVCSSTFNGPFA 1113  
Db 1206 NV--SARFWSLDCPRTAEAVNFATLLYKEL-----TAVPYMAKFV 1243  
Qy 1114 GGVVSPQPHSSYYSGMTGPQHPFNRGSGPAPGV--VLLNSLVNDVACE-----K 1162  
Db 1244 -----IFAKMNDPREGLRCYCMTDDKDTLEQHENFVEVAR 1281  
Qy 1163 LKQIEGLDQSMPLPOYCTTIKKANINGRVLAQCNIDELKKNMNFQDWHLFRSTVLEMRN 1222  
Db 1282 SRDIEVLEGLSL-----FAELSG-----NLVPVKAAQQRSFHQSFRENRLAM-- 1325  
Qy 1223 AESHVVPEDPRFLSESSGPPHGEPPARRASHNE-----LPHTELS----- 1263  
Db 1326 -----PVKVRDSSREPGLSFLRKAMKYEDTQHILCHLNIITWPPCAKSGAEDRR 1376  
Qy 1264 -SQPTYNLNFSFELNLTGLDEGAPR-----HSNLSWQSQTRTP--SLSLNSQD 1311  
Db 1377 RTPTPLALRYSILSESTPGSLSGTEQAEMKMAVISEHLGLSWAELARELOFSVEDIN--- 1433  
Qy 1312 SSIEISKLTQVQAEYDAVREYIAQMSQL-----EGGPGSTTISGRSSPHSTYMGOSSS 1367  
Db 1434 -----RIRVENPNSLLEQSVALLNLWIREG-----QNMENLYTALQSID 1475  
Qy 1368 GGSIHSLNEQKGDSPKPDGRKSFIMKRGDVIDYSSSGVSTNDASPLDPITEDE-- 1425  
Db 1476 RGEIVNMLEGSGRQSRNLKPD-----RRHTDRDYSLSQMNQYSSL-----QDELL 1522  
Qy 1426 -----KSDQSGSKL-----LPKKSERSLSFQTDULKGLSGRLYQKLPSPDE 1467  
Db 1523 SPASLGALSPSLRADYQSEKLPKSKRQDDATGAGQ--DSENEVSLVSGHQRGQARITHS- 1582  
Qy 1468 DE---SGTESDNTPLKDDKDKAEKGVKVERVPSPEHSAREPIRTFIKAEYLSDALDK 1524  
Db 1583 SSLECSKAEDSDAT-----GHEWLEKALSEEPGRPE-----LGSLELVEDDTVDS 1628  
Qy 1525 KDSS-----DSGVRSSESSPNHSLHNEVADDSQLEKANLIBLEDSDSHGKRGIPHSL 1576  
Db 1629 DATNGLDLDLLEQEBQSEKLPKSKRQDDATGAGQ--DSENEVSLVSGHQRGQARITHS- 1686  
Qy 1577 SGLQDPIIARMSICSEDK-----KSPSECSLIASSPEENW 1611  
Db 1687 -----PTVSQVTERSQRLQDWDADGSGIVSYLQDAAQGSW 1721

## RESULT 10

US-09-949-016-5876  
; Sequence 5876, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CU001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5876  
; LENGTH: 1880  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-5876

Query Match 6.3%; Score 555.5; DB 4; Length 1880;  
Best Local Similarity 19.7%; Pred. No. 2.8e-38;  
Matches 374; Conservative 286; Mismatches 581; Gaps 72;

Qy 21 LKALLEKCKVDNERNECGOTPLMIAAECGNLEIVKELIKNGANCNLELDNWTALISASK 80  
Db 94 VRELNVYGANVNAQSGKGFPLVYAAENHLEVKFLENGANQNVADEDTPLAVALQ 153  
Qy 81 EGHVHIIVEELLKCGV-----NLEHRDMGGWTALMWA 111  
Db 154 QGHENVVAHLINYGTKGVRLPALHIAARNDDTRTAALLQNDPNPDLVLSKTGFTPLHIA 213  
Qy 112 CYKRTDVVELLSHGANSVT-----GLQY-----SV 139  
Db 214 AHYENLVNAQLLNRGASVNFTPQNGITPLHIASSRRGNVIMVRLLDGGAQIETKDEL 273  
Qy 140 YPIIWAAGRHADIVHLLLONGAKVNCSDKYGTTPLVMAARKHLECVKILLAMGADV- 198  
Db 274 TPLHCAARNGHVRSEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCCVRLQLQDAEIDD 333  
Qy 199 -----OEGA-----NSMTALIIVAVKGGYTSQVKEILKR 226  
Db 334 ITLDHLTPLHVAACHGHRVAKVLLDKGAKPNSRALNGFTPLHIAACKNHRVVMELLKT 393  
Qy 227 NPVNLTKDGNATMTASKEGTEIVODLLDAGTYNIPDRSGDTVLIGVRGHVIV 286  
Db 394 GASIDAVTESGLTPLHVASFMGHLPIVKNILQRGASPNVSVKVEYETPLHMAARAGTEVA 453  
Qy 287 RALLOKQYADIDIRQDNKTALYMAVEKGNATWVEDILQCPDTEICTKOGTPLIKATKM 346  
Db 454 KYLLONKAKVNAKQDDTPLHCAARIGHTNMVKKLLNANPNLAATTAGHTPLHIAARE 513  
Qy 347 RNIEVBELLDKGAKVSAVDKGDTPHIAIRGSRKLAELLLRNPKDGRLLYRPNKAGE 406  
Db 514 GHVETVLALLEKEASQACMTCKGFTPLHVAAYKGVKVAEELLE--RDA-----HPNAAKG 567  
Qy 407 ---TPYINIDCSHQKSIUQTIFGARHLSPETDGMILGY-DLYSSALADILSEPTWQPIC 462  
Db 568 NGTLPLHVAVHNNLDIVKLLPRGSGSPHSPAWN--GYTPLHIAAKQN-----QVEVA 618  
Qy 463 VGLVAQWGSCKSFLKKLEDEMKTFAQQOIEPLFPQFSMLIVFLTLCCGG---LGLLFA 518  
Db 619 RSLLOYGGSANASV-----QGVTPH-----LAQEGHAENWALLLS 656  
Qy 519 FTVHPNLGIASVLSFLALLYIFFIYVYFGRREGESWNWAVLSTRLARHIGYLELLK- 577  
Db 657 KOANGNLGKSGLTPLHLV-----AQEG-----HVPVADVLKIH 690  
Qy 578 -----LMFVNPPPELPEQTTKALPVRPLF--TDYNRLLSSVG-----GETSLA 616  
Db 691 GVMVDATTRMGYTPHVAHYSHYGNIKLVKFLHQADYNAKTKLGYSPHLQAQQGHTDIV 750  
Qy 617 EMY-----ATLSDACEREPFLA--TRLPFRVFKETDQGGKKWKTKCCLPSFV 662  
Db 751 TLLKNGASNEVSSDGTTPLAIAKRLGYISVTDVLKW-TDET-----SFFV 796  
Qy 663 IFLFIIGCIISGITLAIAPRDPKHL-----TVNAVLIASVVG---LAPVLNCRWTWQ 714  
Db 797 L-----VSKHRMSFPETVDEIL-DVSEDSGEELISFKAERR---- 832  
Qy 715 VLDSLLNSQRKRLHNAASKLHLKSEGFMVKCEVELMARMAKTIDSFQNTQRLVVII 774  
Db 833 --DSRDVDEEKELLDVFPKLDQVVESPAIPRIPC-----AMPETVIVRSEEQASKEY 884

QY 775 DGLDACEQDKVLQMLDTRVRLFS--KGPFAIAFASDP-----HI1IK-----814  
Db 885 DEDSLIPSPATETSDNISVASPVHTGVLSPVMDARGSMRGSHRGLRVVIPPRTCA 944  
QY 815 -----ALNQNLNSVLRDSNINGHDYMN-IVHLPVFL-NSRGLS 851  
Db 945 APTRITRLVPOKLSPTPPLEAEGLASRIIALGPTGAQPLSPVIVEIPHPFASHGRG-- 1002  
QY 852 NARKFLVTSATNGDVPSCDSTTGIEDADRRVQNSLGEMTKGSKTALNRRDTPYRRQWQ 911  
Db 1003 -DRELVLVRSNGSV-----WKEHRSYGBSYLDQILN-GWDEELGSLBLEKXKVC 1052  
QY 912 RTITRQMSFDLTXLVLT--EDWFSDISPOTMRLNIVSVTRGLLRANOISFNWRDLAS 968  
Db 1053 RIIT--TDFPLYPVIMSRILCQD-YDIIGPEG-----GSLSKSLVPLVQATPENAVTK 1102  
QY 969 WINLTQWMPVRTSWLILYLEETEGIPDQMTLK-----TIVERISK----- 1008  
Db 1103 RVKLAHQ-----AGVPDELVTKLGNQATFSPVTVPEPRRKHRRPIGLR 1148  
QY 1009 -NIPTTKDVEPLEIDGDIRNFVF-----LSSRTPVLVARDVKVFLPCTV 1053  
Db 1149 IPLPPSWTDNPRDSGEGDITSLRLCSVIGTDOAQWEDITGTTKLVIYANECANF--TT 1205  
QY 1054 NLDPKLREIADVRAAREQISIGLAYPPLPLHEGPPRAPSGVSPSVCSTSFNGPPA 1113  
Db 1206 NV--SARFWSLSDCPTAEAVNFATLLYKEL-----TAVPYMAKVF 1243  
QY 1114 GGVSVPQPHSSYYSGMTGPHFPYNGSGPAGPV--VLINSLNVDAVCE-----K 1162  
Db 1244 -----IFAKNDPREGRLCYCMTDKDKVKTLEOHENFVEVAR 1281  
QY 1163 LKQIEGLDQMLPOYCTTIKANINGRVLACQNIIDELKEMNMFQDWHLFRSTVLEMRN 1222  
Db 1282 SRDIEVLEGNLS-----FAELSG-----NLVPVKAQAQSRSPHQSFRNRLAM-- 1325  
QY 1223 AESHVVPEDPRFLESSESGPAPGEPARASHNE-----LPHTELS----- 1263  
Db 1326 -----PVKVRDSRSRPPGSGSLSPFLRKAMKYEDTOHLCHLNLTMPPCAKGSGAEDRR 1376  
QY 1264 -SOTPYTLNPSFEELNTGLDECAPR-----HSLMSQSQSRTTTP-SLSSLSQD 1311  
Db 1377 RPTPLALRYSLSESTPPGSLSTEQAEKMWAVISEHLGLSWAELAREIQFSVEDIN--- 1433  
QY 1312 SSIEISKLTKVQAEYRDAYREYIAQMSQL-----EGGPGSTTISGRSSPHSTYMGQSSS 1367  
Db 1434 -----RIRVENPNSLLEQSVALLNLWIREG-----QANMENLYTALQSID 1475  
QY 1368 GGSIHSLNLEQKQKDEPKDQKSLMKRGDVIDYSSSGVSTNDASPLDPITEDE-- 1425  
Db 1476 RGEIVNMLEGGSGRQSNLXPD-----RRHTRDYSLSPSQMGYSYL-----QDELL 1522  
QY 1426 -----KSPQSGSKL-----LPGKKSRSRSLFOTDLKLKGLGSLRYOKLPDSB 1467  
Db 1523 SPASLGCALSSPRAQYNEVAVLDAIPLAATEHDTMLEMSQWVMSAGLPSLVTAE 1582  
QY 1468 DE---SGTEESDNTPLKDKKARBAEKVERVPKSPHSAEPIRTPIKAKEYLSDALDK 1524  
Db 1583 SSLECSKAEDSDAT-----CHEWKELGALSEEPRGPE-----LGSLELVEDDVTDS 1628  
QY 1525 KUOS-----DSGVRSESSPNHSLNHEVADDSQLEKANLIELEDDSHSGKRGPHSL 1576  
Db 1629 DATNGLIDLEQEGORSEKLPFGSKRQDDATGAGQ-DSENEVSLVSGHQRGQARITHS- 1686  
QY 1577 SGLQDPIIARMSICSDK-----KSPSECSLIASSPEENW 1611  
Db 1687 -----PTVSQVTERSQDRLOQWDADCSIVSYLQDAAQGSW 1721

RESULT 11

US-09-949-016-6965

; Sequence 6965, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTIER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6965

; LENGTH: 1881

; TYPE: PRT

; ORGANISM: Human

; US-09-949-016-6965

Query Match

Best Local Similarity 6.3%; Score 555.5; DB 4; Length 1881;

Matches 374; Conservative 286; Mismatches 659; Indels 581; Gaps 72;

QY 21 LKALLEKCKVDNERNECGOTPLMIAAEQGNLEIVKELIKNGANCULEDNWTLALISASK 80

Db 94 VRELNYGANVAQSQGKTPLYMAAQENHLEVYKFLLENGANQNVADEGTFPLAVALQ 153

QY 81 EGHVHIVBELKCGV-----NLEHRDMGSGWTALMWA 111

Db 154 QGHENVVAHLINVTGKVRPLPALHIAARNDTRTAAVLLQNDPNDVLSKTGTFPLHIA 213

QY 112 CYKGRDVTVELLSHGANSPT-----GLQY-----SV 139

Db 214 AHYENLVNAQLLNRGASVNFPTQNGITPLHIASRRGNVIMVRLLLDRGAQIETKTDEL 273

QY 140 YPIIWAAGHADIIVHLLQNGAKVNCSDKYTTPLVWAARKGHLECVKHLAMGADV- 198

Db 274 TPLHCAARHVRISILLDHGAPIQAKTNGLSPIHMAAQGDHLCVRLLLQYDAEIDD 333

QY 199 -----QEGA-----NSMTALIVAVKGGVTSQVKILKR 226

Db 334 ITLDHLTPLHVAACHGHRVAKVLLDKGAKPNRSLNGTPLHIACKKHVRVWELLKT 393

QY 227 NPNVNLTRDNGNTALMIASKEGHEITVQDLDBAGTYVNIPIRSGDTVLICAVRGHVEIV 286

Db 394 GASIDAVTESGLTPLHVSFMGHLPIVKNLQRGASPNVSNVKVETPLHMAARAGTEVA 453

QY 287 RALLQYADIDIRGQONKTALYWAKEGNATWVRDILQCNPDTEICTKQGETPLIKATKM 346

Db 454 KYLLQKAKVNAKAKDDQTPHCAARIGHTNNVYKLLLENNANPNLATTAGHTPLHIAARE 513

QY 347 RNIEVVLLLDKGAQVSAVDKGDTPHIAIRGRSEKLAELLLRNPKDGRLLYRPNKAGE 406

Db 514 GHVETVLLALLEKASQACMTKKGFTPLHVAAKYKVRVAELLE--RDA-----HPNAAKG 567

QY 407 ---TPYNIHCSHQKSLTQIFGARHLSPTETDQDMLGY-DLYSSALADILSEPTMOPPIC 462

Db 568 NGITPLHVAVHNNLDIVKLLPRGSGSPHSPAWN--GYTPLHIAAKQN-----QVEVA 618

QY 463 VGLYQWQSGKFLKLEDEMKTFAQOIEPLPQSPWLIIVFTLLLCGG-----LGILFA 518

Db 619 RSLQYGGSANAEV-----QGVTPH-----LAAQEGHAEMWALLLS 656

QY 519 FTVHPNLGIATVSLFLALYVIFFIYFGRREGESNNWAVLSLRLARHIGVLELLK- 577

Db 657 KQANGNLGNKSGLTPLHLV-----AQEG-----HVPVADVLIK 690

QY 578 -----LMFVNPPELPEQTTKALPVRF--TDYNRLSSVG-----GETSLA 616

Db 691 GWMVDATTRMGVTPHVAASHYGNIKLVKFLQHQADVNAKTKLGYSPLHQAQOQHTDIV 750

Qy	617	EMI-----ATLSDACREERGFLA--TRLFRVFKTEDTQGGKKWKTKCCLPSFV	662
Db	751	TLKNGASPNVSSDGTTLTAIAKRLGYISVTDVLKVV--TDET-----SFF	796
Qy	663	IFLFIIGCIISGITLLAIFVDPKHL-----TVNAVLISIASVVG---LAFVLNCRTWQ	714
Db	797	L-----VSDKRMSPFETVDEIL--DVSEDEGBELISFKARR-----	832
Qy	715	VLDLSLNSQRKLHNAASKLHKSEGMVKLVKCEVELMARMAKTIDSFQNTQRLVVII	774
Db	833	--DSRDVEEKELLDFVKLDQVVESPAIPRIPC-----AMPETVVIRSEGEQASKEY	884
Qy	775	DGLDACEODKVLQMLDTRVVLFS--KGPFTAIFASDP-----HIHK-----	814
Db	885	DEDSLIPSPATETSDNISPVASPVHTGFLVFMVDARGGSMGRSRRNGRUVVIPPRTCA	944
Qy	815	-----AINONLSNVLRRDSNINGHDYMRN--IVHLPVFL--NSRGLS	851
Db	945	APTRITCRLVKPQLSTPPPLAEBEGLASRIIALGPTGAQFLSPVIVEIHPFASHGRG--	1002
Qy	852	NARKFLVTSATNGDVPSCDGTGTIGQEDADRVSQNSLGEMTKLGSKTALNRDRDYRRRQMQ	911
Db	1003	-DRELVLRGENSV-----WKEHRSRYGESYLDQILN-GWDEELGSLELEKKRVC	1052
Qy	912	RTITROMSFDTKLKLVIT--EDMFSDISPQTMRELLNIVSVTCELLRANQISFNWDELAS	968
Db	1053	RIIT--TDFPLYFVIMSRLCOD--YDIGPEG-----GSLGKGLVPLVQATFPENAVTK	1102
Qy	969	WINTEQWPVRTSWLILYLETETGIPQOMTLK-----TIYERISK-----	1008
Db	1103	RVKLAQO-----AQPVPDELVTKLGNQATSPVIVTEPRRKRHRPIGLR	1148
Qy	1009	-NIPTTKDVBPLEIDGINFEVF-----LSSRTPVLVARDVKVFLPCTV	1053
Db	1149	IFLPPSTWTDNPRDSGEDTTSLRLCSVIGGTQQAOWEDITGTTKLYANECANF--TT	1205
Qy	1054	NLPDKLREIADVRAAREQISIGGLAYPPLPHLEGPPRSPGVSQPPSVCSSTSFNGPFA	1113
Db	1206	NV--SARFVUSDCPRTAEAVNFATLLYKEL-----TAVPYMAKFV	1243
Qy	1114	GGVVSPOPHSSYYSGMTGPQHPFNRRSGPAGPV--VLLNSLNVDVACE-----K	1162
Db	1244	-----IFAKMNDPREGLRCYCMTDDKVDKTLQHENFVEVAR	1281
Qy	1163	LKQIEGLDQMLPOYCTTIKANINGRVLACNIDELKEMNNMFGDWHLFRSTVLEMRN	1222
Db	1282	SRDIEVLEGMSL-----FAELSG-----NLVPVKKAAQORSFHQSFRENRLAM--	1325
Qy	1223	AESHVWPEDPREFLESSESGPAPGEPARRASHNE-----LPHTELS-----	1263
Db	1326	-----PVKVRDSREPGGSLFURKANKYEDTOHILCHLNITMPPCAKSGGAEDRR	1376
Qy	1264	-SOTPYTLNFSFBEINTLGLDEGAPR-----HSNLWSQOTRRTP--SLSSLSNQD	1311
Db	1377	RTPTPLALRYSIILESPTGSLSGTEQAEMKMAVISEHGLSWARELQFVEDIN--	1433
Qy	1312	SSIEISKLTQVQAYRDAYREYIAQMSQL-----EGGPGGTTISGRSSPHSTYMGQSSS	1367
Db	1434	-----RIRVENPNSLLEOSVALLNLWIREG-----QNaNMENLYALQOSID	1475
Qy	1368	GGSIHNSLEQKGDSEPKDDGRKSFPMKRGDVIDYSSGVSNTDASPLDPTITEDE--	1425
Db	1476	RGEIVNMLEGSGRQSNLKPD-----RRHTORDYSLFSPQMWGYSSL-----QDELL	1522
Qy	1426	-----KSDQSGSKL-----LPGKKSRSRSLFQDTLKLKXGSLRYQKLPSPDE	1467
Db	1523	SPASLGCALSSPLRADQYWNVEAVLDAIPLAAATEHDTMLSMSQMVWSAGLTPSLVTAED	1582
Qy	1468	DE---SGTESDWTPLKDDKDKRAEKVERVPKSPHSEAPRTPTFKAEYLSDALLDK	1524
Db	1583	SSLECSKAEDSAT-----GHEWKLEGALSEEPGRPE-----LGSLELVEDDTVDS	1628

```

Qy 1525 KDSS-----DSGVSRSSSPNHSLHNEVADDSOLEKANLIELEDDSHGKRGIPHSL 1576
      ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1629 DATNGLIDLLEQEGQSEKLPCKSRQDDATGAGQ-DSNEVSLVSGHQRGQARITHS- 1686
      ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Qy 1577 SGLQDPITIAIRMSICSEDK-----KSPSECSLIASPEENW 1611
      ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1687 -----PTVSQVTERSQDLQDWDADGSGIVSVLQDAAGQSW 1721
      ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 12
US-09-949-016-7659
; Sequence 7659, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7659
; LENGTH: 2753
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7659

```

Query Match 6.28; Score 549; DB 4; Length 2753;

```
QY 553 ESNWAWVLSTRLARHIGYLELL--LKLM----- 579
Db 662 TTANGNTALA--IAKRLGYISVVDTLKVVTVEVTTTTTITEKHKLNPETMTVEVLVDSD 719
QY 580 -----FVNPPPEOTTKALP--VREL-FTDYNRLSSVGGETSLAEMIATLSD 624
Db 720 BEGDDTMTGDGGEYLPEDLKEGLDSDLPSSQFLDGMVYLRYSLEGRSDSDRSHTLH 779
QY 625 AC-----ERFEGFATRLFRVFKTEDTOGKKW-----KKTCLPLS 660
Db 780 ASYLRDSAVMDDSVIPSHQVSTLAK-----EAERNYSYRLSMGTENLONVALSSSPIHS 833
QY 661 FVIFLII-----GCIISGITLAI FRVDPKHLTVNAVLIASVGLAFVLCRTW 712
Db 834 GLVSPMVDARGAMRGCRHNGRLII--IPRCKTAPT-----RVTCL-- 875
QY 713 WQVLDLLNSQKRL-----HNAASKLHLKXSEG-----FMKVLKCEV 750
Db 876 -----VKRHLATMPMVEGEGLASRLIEVGSQAFLGPVIVEIPHFAALRGKER 926
QY 751 ELWARMAKTIDSTON-----QTRLVVIIDGLDACEQDKVQLQMLDTRVRLFSKGPFIAP 805
Db 927 ELVWLRSENGDSWKEHFCVDTEDELNEILNGMD-----EVLDSPEDEKKRICRIIT 978
QY 806 ASDPH--IIIKAINONLNSVLRDSNNGHDYMRNIVHLPVFLNSRGLSNARKPLVTSATN 863
Db 979 RDPQYFAVVSRIKQ-----DSNLIGPE-----GGVLSSTVVPQVQVAFPEGALT 1023
QY 864 GDVPCSDTTGIOEDADRRYSQNSLGBMTK--LGSKTALNRDFTY--RRQMQORTIFROMS 919
Db 1024 KRI-----RVGLQ-----AQPHSELVKILGNKATFPVITLPERRKFKHPITWTIP 1072
QY 920 FOLTKLVLTEDWFSDISPOTMRRLNIVSVTGRLLRANOISFNWDRKASWINLT---EQW 976
Db 1073 VPKASSDVMNGFGDAP--TLRL--CSITGTTPA-----QWEDITGTTPLTFVNECV 1123
QY 977 PYRTS-----WLILYLEETEGIPDQMTLTIYERISKNIPTTKVDPELLEIDGDIRNFEV 1031
Db 1124 SFTTNVSARFWLI-----DCRQIQESTFASQVYREIICVPY-----MAKFVY 1166
QY 1032 FLSSRTPVLVARDVKVFLPCTVNLDPKL-----REIIDVRAAREQISIGGLAYPLPL 1085
Db 1167 FAKSHDP-----IEARLACFCMTDDKVKTLEQOENFAEVARSRD-----VEV 1209
QY 1086 HEGPPRAPSGYQPPVSCVSTSNGP-----FAGGVVSQPHSSYSGMTGPOHPF-- 1136
Db 1210 LEGKPIYVDFGNLVLTKSTRYDSGTLRTTIYEGAKIHERPGASSYLQL---KHFFAD 1266
QY 1137 -YNRSGPARGPVV---LLNSLVNDVACEKLQIEGLDQSMPLPOYCTTIKKANINGRVL 1191
Db 1267 LYKDDTESETSVLKSHLVNEVPVLASPDLLSEVSEMQDLTKM--TAILTDDVSDKA- 1323
QY 1192 AQCNIDELKEMMNFQDWHLPSTVLEMRNABSHVVVPDPRFLSE--SSSGPAPHCEPAR 1250
Db 1324 GSIKVVELKAAEEEGE-----PFEIVER-----VKDELEKVNELISGICTRDESV 1372
QY 1251 RASHNELPHTLESSQPTTYTNFSEBELNTGLDEGAPRHSNLSWQSQTRTPPSLSSLSNQ 1310
Db 1373 QSSRSERGLVE--EEMVIVSDEIEEAR-----QKAPLEITEYPCVEVIDKEIKGVKEK 1425
QY 1311 DSSIEISKLTDKQVARYRDARYIAQMSQLEGPGSTTISGRSSPHSTHYMGQSSSGGS 1370
Db 1426 DSTGLNYLITDOLLNT--CVPLPKELQTVQDKAGKCEALAVGRSS----- 1469
QY 1371 IHSNLBOEKGKQSEP-----KPDGRKSFMLKMGDVIDYSSSGSVSTNDASPLDPI 1420
Db 1470 -----EKEGKOLPPDETOSTQKHPSIGIKKPVRRK-----LKEK 1505
QY 1421 TBEDEKDSQSKLPGKKSRSLSLFTQDLKLKGSGLRYQKLPSPDESGTEESDNTPL 1480
Db 1506 QKQKEGLQASABKAEKLGKSSSESLGE-----DPGLAPEPLPTVKATS-----PL 1551
```

```
QY 1481 LKDDKDKRKAEGKVERVPKSP-----HSAEPIRTFIKAEYLSDALLDKKDSSDSGVRSS 1535
Db 1552 IBETPIGSIKDKVKALQKRVEDBQGRSKLPIR--VKGEDVPKTKTHRPHPAASPSLKS 1609
QY 1536 E-----SSPNHLSHNEVADDSQLEKANILEEDDSHSGKRGIPHSLSGLQDPIIARMSI 1589
Db 1610 ERHAPGSPKTERHSTLSSSAKTERHPPVS-----PSSKTEKHSVP---SPS 1654
QY 1590 CSBEDKKSPECSLIASSPENWPAQOKAYNLNRTPTSTVTLNNSAPANRANQNDFEMEGI 1649
Db 1655 AKTERHSPASS--SKTEKHSVP-----SPSTKTERHSPVSTTERH---PPV 1698
QY 1650 RETSOVILRPSSSPNPTTIONENLK--SMTHKRSQSSSYTRLSKDDP 1694
Db 1699 SPSGKTKDRKPPVSPSGRTEKHPVSPGRTEKRLPVPSPGRTDKHP 1744

RESULT 13
US-09-949-016-7660
; Sequence 7660, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7660
; LENGTH: 2753
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7660
```

```
Query Match 6.2%; Score 549; DB 4; Length 2753;
Best Local Similarity 21.2%; Pred. No. 2.2e-37;
Matches 392; Conservative 284; Mismatches 710; Indels 460; Gaps 72;
```

```
QY 21 LKALLEKCKDVERNECGQTPLMIAABQGNLIVKELIKNGANCNLEDLDNWTALISASK 80
Db 187 VELLLERGAPLLARTKNGLSPLHMAAQGDHVECVKHLQHKAPVDDVLDLTALHVAH 246
QY 81 EGHVHIVBELLLKGVNLEHRDMGGWTALMWACVKGRTDVVELLSHGAN--PSVTGLQYSV 139
Db 247 CGHYRVTKLLDKRANPNARALNGFTPLHIACKQRIKVMELLYKYGASIQAIT--ESGL 304
QY 140 YPIJAAAGHGHADIHVHLLQNGAKVNCSDKYGTTPLVAAARKGHLECVKHLAMGADVQ 199
Db 305 TPIHVAAPFMHLNIVLLLLQNGASPDVTNIRGTALHMAARAGQVEVRLRNGALVDA 364
QY 200 EGANSMTALIVAKGYTQSVKEILKRNPNVNLTDKGNLTALMIASKEGHEITEIVQDLDA 259
Db 365 RAREEQTPHIASRLGKTEIVQLLLQHMAHPDAAATNGYTPHLHISAREQVDVASVLEA 424
QY 260 GYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRGQDNKTALTVAWEKGNATMV 319
Db 425 GAAHSLATKKGFTPLHVAAYKYGSLDVAKLLQRRRAAADSAGKNGLTPHVAHYDNQKVA 484
QY 320 RILQCNPTETCTKDGTEPLIKATKORNIIEVELLDKGAKVASVDKGDTPPLHIAIRG 379
Db 485 LLLLEKGAAPHATKNGYTPHLHIAKQNMQASTLUNYGAETNIVTKQGVTPHLHIAEQ 544
QY 380 RSRKLAELLLRNPKQGRLLYRPNKAGETPYNIDCSHQK----SILTOIFGARHLSPTETD 435
Db 545 GHTDMVTLLL---DKGANIHMSTKSLTSLHLLAAQEDKVNADILTK--HGADQDAHTK-- 598
```





Db 253 LNKGAAVDFTA-RNGITPLHVASKRGNTNNVKKLLDRGGQIDAKTRDGLTPLHCAARSGH 311  
QY 183 -----HLECVKHLAMGADVDOEGANSMALIV 210  
Db 312 DOVVELLLERGAPLLARTKNGLSPLHMAAQGDHVECVKHLQKAPVDDVTLDTLALHV 371  
QY 211 AVKGGYTOSKEITLKNPNVNLTKDGNALMIASKE----- 247  
Db 372 AHCHGYRVTKLLDKRAHPNARALNGFTPLHIAACKNRIKWNELIVKYGASIQAITESG 431  
QY 248 -----GHTEIVQDLDAGTYVNIIDRSQDVTLIGAVRGHVEIVRALIQKYADID 297  
Db 432 LTPHVAFAFMHNLNVLILLQNGASPDVTNIRGETALHMAARAGQVYVRCLLRNGALVD 491  
QY 298 IRQDNKTALYWAVEKGNATWDRDIILOCPDPTICTKQGETPLIKATKRNIEVEVLDD 357  
Db 492 ARAREEQTPHIAISRLGKTEIVQLLQHMAHPDAATTNGYTPLHISAREQQVDVASVLE 551  
QY 358 KGAKVSAVDKGDTPHIAIRGRSRKLAELLLR-----NPKD 394  
Db 552 AGAHSATYKGTPLHVAKYGSLVAKLLQRRRAADSAGKNGLTPLHVAHYDNQKV 611  
QY 395 GRLL-----YRPNKAGETPNIDC-SHOKSILTQI--FGAR-----HLSPT 432  
Db 612 ALLLEKASPHATAKNGYTPLHIAAKNQMQIASTLLNVGAETNIVTKQGVTPHLASQ 671  
QY 433 ETDGDM-----GYDLYSALADIISEPTMQPPICVGLYAQWGSKSLKKLEDEM-- 484  
Db 672 EGHDMVTLLLDKGANIHMSKSLGTS-----LHLAAQ-----EDKVV 710  
QY 485 -----KTEAGQOIEPLFQPSWLVFLTLCL--GGGLL--FATVHPNLGIAVLSFLAL 536  
Db 711 ADILTKHGADQDHTKGLGYTPLV-----ACHYGNVMNVLKQGANNAKTKNGYTPL 765  
QY 537 -----LYIFPIVYFGRREGESNMWAVLSTRLARHIGYLELL--LKL----- 579  
Db 766 HQAAQOQHTHINVLQHGAKPNATTANGTALA--IAKELGYISVVDTLKVTEEVTT 823  
QY 580 -----FVNPPPLPQTKALP-VRFL-F 600  
Db 824 TTTITEKHLNVPEITVELVDSDEGDDTMTGDGGEYLRPELDELGDSDLSPSSQFLDG 883  
QY 601 TDYRLSSVGET-SL-----AEMIATLSDACERE 629  
Db 884 MNLYRLSLEGGSRDSLRSFSSDRSHTLASHAYLRDSAVMDDSVIIPSHQVSTLAKAERN 943  
QY 630 FGLATRLFRVFKTEDTOGKKWKTCCLPSFVIFLF-----IIGCIISGITLLAIFRV 683  
Db 944 SYPLS-----WGTENLDNLVALSSSPIHSGFLVIFWVDARGGAMRGCHNGLRII----I 993  
QY 684 DPKHLTVNAVLSIASVWGLAFVLNCRTWQVLDLSLLNSQRKL-----HNAASKL 734  
Db 994 PPRKCTAPT-----RVTCRL-----VKRHLATWPMVEGEGLASRL 1030  
QY 735 HKLKSEG-----EMKVLKCEVELMARMAKTIDSFQNTQ-----QRLVVIIDG 776  
Db 1031 IEVGPSSGAOFLGPIVIEIHFHAALRGKRELVLVRSENGDSWKHECFDYTEDNELNGLNG 1090  
QY 777 LDACEQDKVLQMLDTRVRLFSKGFPIAIFASDPH--IIIKAINQNLNSVLNRDSNINGHY 834  
Db 1091 MD-----EVLDSPEDEKKEIKRIITRDPQYFAVVSRIKQ-----DSNLIGPE- 1134  
QY 835 MRNIVHLPVLNRSRGLSNARKFLVTSATNGDVPSCOTTGIEQDADRRVQNSLGEWTK-- 892  
Db 1135 -----GGVLSSTVVPQVQAVPEGALTAKRI-----RVGLQ-----AQPMSHSELVKKI 1176  
QY 893 LGSKTALNRDITY--RRRQQRITITQMSFDLTKLAVTEWDSIDISQPTMRLLNIVSVT 950  
Db 1177 LGNKATFPIVLEPRRPHKFDITWIPVFKASSDVMLNGFGGDAP-TLRL--CSIT 1232  
QY 951 GRLLRANQISFNWDRLASMINIT--EQWPYRTS-----WLILYBETEGIPDMQTLKT- 1001  
Db 1233 GGTTPA-----QWEDITGITPLTFVNECVSFTTNVSARFWLI-----DCRQIQESVTFASQ 1283

QY 1002 IYERI-----SKNIPTTKDVEPLL-----EIDGDIRNFVFLSSRTPVLVARD 1044  
Db 1284 VYREIICVYMAKFFVFAKSHDPIEARLCFCMTDDDKVKTLEQENF--AEVARSRD 1339  
QY 1045 VKV-----FLPCTVNLDP-----KLREIITAD-----VRAAR 1070  
Db 1340 VEVLKGPPIYVDCFGNLVPLTKSGQHHIFSPFAFKNRNLPLFKVVRDITQEPGCRISPMK 1399  
QY 1071 EQISIGGLAYP-----FLPLHEGPPRAPSGYQSPSCSSTSFNGPAGGVVSPQPHS 1123  
Db 1400 EPKSTRGLVHQAICNLNITLPIY--TKSESQOEQEEIDMTSEKN-----DETES 1448  
QY 1124 SYSGMTGQHPFPYNRGSPAPGVPVLLNSLAVDAVEKLAQIEGLDQSLMPOYCTTIKK 1183  
Db 1449 TETSVL--KSH-----LVNEVPVLASPDLLSEVSEMKQDLIKM--TAILT 1489  
QY 1184 ANINGRVLAQCNIDELKEMNMNFGDWHLFRSTVLEMRNAESHVVPEDPRFLSE-SSSGP 1242  
Db 1490 TVUSDKA-GSIKVKELVKAABEPGE---PFEIVER-----VKEDLEKVNELIRSGT 1537  
QY 1243 APHGEPPARASHNELPHTELSQTPYTLNFSFEELNTLGLDEGAPRHSNLSWOSQTRTP 1302  
Db 1538 CTRDESSVQSSRSRERGLVE--BEWIVSDBEIEEAR-----OKAPLEITEYPCVEVRIDK 1590  
QY 1303 SLSLSNSQSSSIEISKLTQVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYM 1362  
Db 1591 ETKGKVEKQDSTGLVNYLTDDLANT--CVPLPKEQLQTVODKAGKKEALAVGRSS----- 1642  
QY 1363 QGSSSGGSIHSLNBOEKHKDSEP-----KPDGGRKSLMKRGDVIDYSSSGVSTN 1412  
Db 1643 -----EKEGKDIIPDETQSTQKHKPSLGIKKPVRRK----- 1674  
QY 1413 DASPLDPIITEDEKSDQSGSKLLPQKKSSERSLFTDLKLGSGLYRYKQLPSDEDESGT 1472  
Db 1675 ----LKEQKQKEGLQASAEKAEKAKKSGSESLGE-----DPGLAPEPLPTVKATS-- 1722  
QY 1473 ERSNDTPLKDDKDKAEKGVKERVKSPKPE-----HSAEPIRTFIKAEVYLSDALLOKDS 1527  
Db 1723 -----PLIETPIGSKDKVAKLQKRVEDQKGRSKLPFR--VKGEDVPKKTTHRPH 1774  
QY 1528 SDGVRASSE-----SSPNHSLHNEVADDSQLEKANLIELEDDSHSGKRGIPHSLGLQD 1581  
Db 1775 AASPSLKSRRHAPGSPKTERHSTLSAKTERHPVS-----PSSKTEKHS 1822  
QY 1582 PIARMSICEPKSPSECSLIASSPEENWPACQKAYNLNRTPTVTYTLNNSAPANRANQ 1641  
Db 1823 PV---SPSAKTERHSPASS--SKTEKSPV-----SPSTKTERHSPVSTKTER 1867  
QY 1642 NDEMEGIRETSQVILRPSSSPNPTTIQENLX-SMTHKRSQBSYTRLSKDDP 1694  
Db 1868 H----PPVSPSGKTKDRPPVSPSGRTEKHPVSPGRTEKRLPVSPSGRTDKHQP 1917

## RESULT 15

US-09-172-977-4  
; Sequence 4, Application US/09172977  
; Patent No. 5989863  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
; FILE REFERENCE: PF-0615 US  
; CURRENT APPLICATION NUMBER: US/09/172,977  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 1839  
; TYPE: PRT  
; ORGANISM: Homo sapiens

;  
;  
US-09-172-977-4

Query Match 6.0%; Score 530.5; DB 2; Length 1839;  
Best Local Similarity 20.2%; Pred. No. 4.2e-36;  
Matches 426; Conservative 263; Mismatches 654; Indels 763; Gaps 81;

QY 14 EENIPALKALLEKCKDVDERNECGOTPLMIAAOGNLEIVKELIKNGANCNLDELNDWT 73  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 73 KEHVGVLVQELLGRGSSVSATKGNLTALHIASLAGOAEVKVLVKEGANINAGSQNGFT 132  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 74 ALISAKEGHVHIVEELLKGVN-----LEHRDMG--- 103  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 133 PLYMAAQENHDVVKYLLENGANQSTATEDGFPLVALOOQHNOAVAILLENDTKGVR 192  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 104 -----GWITALMWACYKGRDVTUVELL 123  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 193 LPALHIAARKDDTKSAALLONDHADVQSKMVNRRTESGFTPLHIAHYGNVNVA TL 252  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 124 LSHGANPSVTGLQYSVPYPIIWAAGRHADIVHLILONGAKVNSDKYTGTTPLVMAARKG- 182  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 253 LNREGAAVDFTA-RNGITPLHAVSKRGNTNNVKLLDRGGQIDAKTRDGLTPLHCAARS GH 311  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 183 -----HLEVCVKKHLAMADVVDQGANSMTALTIV 210  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 312 DQVVELLERGAPLLARTKNGLSPLHMAAQGDHVECVKHLLOHKAPVDDVTLDYLTAHV 371  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 211 AVKGGVYQSVEIKENPNVNLTDKCNTALMTASK----- 247  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 372 AAHCGRYRVTKLDDRANPNARALNGFTPLHACKQRICKWELLYKVYGASTQAITE SG 431  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 248 -----GHTEIVQDLLDAGTYVNIPIORSGDTVLIGVRGHEIVEIRALLQKYADID 297  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 432 LTPIHVAAFMGHLNIVLLLIQNGASPDTVNIARGETALHMAARAGQEVVRCLLRGALVD 491  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 298 IREQDNKTALYWAVEBKNATMPWDILQCNPDTETICTKDGETPLIKATKMENIEVELLLD 357  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 492 ARARESQTPLHASRLGKTEIVOLLQHMHPDAATTNGYTPLHISAREGOVDVASILLE 551  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 358 KGAKVSAVDKDGDTPLHIAIRGRSRKLAELLLR-----NPKD 394  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 552 AGAAHSLATKKGFTPLHVAAYKGYSLDVAKLLQRRRAADSAGKNGLTPLHVAHYDNQKV 611  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 395 GRILL-----YRPNKAGETPPNYDC-SHOKSILTQI--FGAR-----HLSP T 432  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 612 ALLLLEKGA SPHATAKNGYTPLHIAAKNQMQIATSTLLNYGAETNIVTKQGVTPHLAS Q 671  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 433 ETDGDML-----GYDLYSALADIILEPTMQPPICVGLYAONGSGKSFLCLKLEDWM-- 484  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 672 EGHDTMVTLLDKGAINHMSKSGLT S-----LHLAQ-----EDKNV 710  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 485 -----KTFGAQQOTEPLFOPSWLIVFLTL LC--GGJGLL-PAFTVHPNLGIAVLSLFAL 536  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 711 ADILTKHGADQDAHTKLGVTPLIV----ACHYGNVMNVNFKLQGANVNAKTNGYTP L 765  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 537 -----LYIPVIYIFGRRGESBNWAWLSTRLARHGYLELL--LKLM----- 579  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 766 HOAAOQGHTHINVLLQHGAKPNATTANGWTALA--IAKGLGISYVVDTLKVVEEVT TT 823  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 580 -----FVNPPPEPQTTKALP-VRF-L-F 600  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 824 TTTITEKHKNVPETWTEVLVDSEBGDTMTGDGGEYLPEDIKELGDDSLSRSSQLDG 883  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 601 TDYNRLSSVGGET-SL-----AEMIATISDACERE 629  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 884 MNLYRSYLEGRSDLSRFS SDRSHTLSHASYLRDSAVMDSDSVVIPHQVSTSLAKEARN 943  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 630 FGLFATRLFRVFXTEDTQGGKKWKTCCLPSFIELF-----IIGCIISGITLLAI FRV 683  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
Db 944 SYRLS-----WGTEINDNALSSSPIHSGFLVIFWVDARGGMARGCRHGLRII --- 993  
|| : || : || : || : || : || : || : || : || : || : || : || : || :  
QY 684 DPKHLTVNAVLTIASVVGVLAFVLCNCTRWQVLDSLNSQRKL-----HNAASKL 734  
|| : || : || : || : || : || : || : || : || : || : || : || : || :

Db	994	PPRKCTAPT-----RVTCRL-----VKRHLATMPMVGEGLASRL	1030
Qy	735	HKLKSEB-----FMKVLKCEVELMARMAKTIDSFTON-----QTRLWVIIDG	776
Db	1031	IEVPGSAQFLGPVIVIEIHPHFAALRKERELVLRSENGDSWKEHFCDYTEDÉLNLNG	1090
Qy	777	LDACEODKQLMDLTVRVLFSKGPFAIFASDPH--IIIKAINQNLSVLNDSNGHNDY	834
Db	1091	MD-----EVLDSPEDLEKKRICRIITRDFQPYFAVVSRIKQ-----DSNLGPE-	1134
Qy	835	MRNIVHLPFLNRSGLSNARKFLVTSATNGDVPSCDTTGIQEDADRVRVQNSLGEWTK--	892
Db	1135	-----GGVLSTVVQPQAVPEPBGALTKRI-----RVGLQ-----AQPMHSEIVVKKI	1176
Qy	893	LGSKTALNRDVTY--RRRQMQRTITROMSFDLTCLKLVTDWFSDISDPQWRRLLNLVSVT	950
Db	1177	LGNKATFSPITVLEPRRRFKHPKITMTI PVPKASSDVMLNGFGGDAP-TLRL--CSIT	1232
Qy	951	GRLLRANQISFNWDLRASWINUT--EQMPYRTS-----WLILYLBETEGIPQDMTKTI	1002
Db	1233	GGTTPA-----QWEDITGTTLPTLVNCECVSFTTNVSARFWLI-----DCRQIOBSVTFASQ	1283
Qy	1003	YERISKNIPTTKDVEPLLEIDGIRNFVFLSRTPVLVAROVKVPCLPCTVNLDPKL---	1059
Db	1284	VYREIICVPY-----MAKFVVVFAKSHDP-----IEARLRCFCMTDDDKVDKT	1324
Qy	1060	---REIIADVRAAREQISIGLAYPPLPLHEGPPRAPSGVSPVPSVCSSTSFNGPAGGV	1116
Db	1325	LEQOENFAEVARSD-----VEVLEKPIYVOCF-----GNL	1356
Qy	1117	V---SPQPHSSYYSGMTGPHFPYNRGSGPAGPVPVLLNSLNVADACEKLQIEGLDOS	1172
Db	1357	VPLTKSGQHIFSPFAFKENRLPLFKVURDTTQEPGGRLSFM-----KEPKSTRGLVHQ	1410
Qy	1173	MLPOYCTTTKKANINGRVLAQCNIID-----ELKEMMNMFQDWHLFRSTVL	1218
Db	1411	-----AICNLANTLPIYTKESDQOEIEIDMT-----	1439
Qy	1219	EMRNAESHVVPEDPRFLSESSGPAHPGEPARRASHNELPHTLELSOTPYTLNFPSEELN	1278
Db	1440	-----SEKNQDEQRIBE-----RLAYIADHLGFSWTELARE---LDTFEQIHL	1481
Qy	1279	TLGLDEGAPRHNLSWQSQ-----TRRTFSLSSLSQSDSIEISKLTDK	1322
Db	1482	QIRIE---NPNSLQDOSQYLLKIWLERDQKATDNLVECLTKINRMDIVHLMETNTEP	1537
Qy	1323	VQAEYRDAVREYIAQMSLEQGGSTTISGR--SSPHSTYYMGQSSSGGSIHNSLBOEKG	1380
Db	1538	LQERISHSYAE-IEQITIDHSEGFSLQEEELCTAQHK-----QKEQOAVS	1582
Qy	1381	KDSEKPEDDGRKSFMLMKRGDVIDYSSGSVSTNDASPLDPITEED-----EK	1426
Db	1583	KESE-----TCDHPPI---VSBEDISGVYSTQDGVPKTE	1614
Qy	1427	SPQSGSKLLPGKKSSERSSLFQDLKLKSGGLRYQKLPSEDED-ESGTEESDNTPLLKDDK	1485
Db	1615	GNSSSTALFPQTHKEQ---VQODF---SG-KWQDLPSESLYEQOEYFVITPGTETSE	1665
Qy	1486	DKAEKQKVRVPKSPHSAEPRTFTIKAEYILSDALLDKDSDSDGVRSESSPNHSLHN	1545
Db	1666	TOKA-----MIVPSSPSKTPE-----	1681
Qy	1546	EVADDSQLEKANLIELEDDSHSGKRGIPHLSGLQDPIIARMSICSE--DKKSPSECSLI	1603
Db	1682	EYSTABEBK---LYLQTFTSBERG-----SPIIQEPPESEHKEESSPRKTSILV	1729
Qy	1604	ASSPENWP-----ACOKAYNLNR-TPSTVTLNNSAPANRANQNPMEGIRE-	1651
Db	1730	IVESADNQETCERLDEDAAFKGDMDPEIPEITVT-----EEYIDSH	1773
Qy	1652	-----TSQVILRPSS-----PNPTTIQENLKSMTWTKRQORSYTRL	1689

Db 1774 GHTVVKVTRKIIIRYVSSSGTEKEEIMVQMPQEPVNIIEGDSYKVIKRVVLKSDTEQ 1833

Qy 1690 SKDPPE 1695

Db 1834 SEDNNE 1839

Search completed: February 5, 2005, 19:36:15  
Job time : 70.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:28:23 ; Search time 141 Seconds  
(without alignments)  
3961.558 Million cell updates/sec

Title: US-10-021-571-4  
Perfect score: 8884  
Sequence: 1 MSVLISQSVINYVEENIPA.....LHAAASSTGFGBERESIL 1715

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pcp.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pcp.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8884	100.0	1715	14	US-10-021-571-4
2	8884	100.0	1715	14	US-10-117-229-2
3	8884	100.0	1715	16	US-10-473-574-26
4	8830	99.4	1771	14	US-10-117-229-4
5	8830	99.4	1771	14	US-10-117-229-11
6	8362.5	94.1	1762	14	US-10-117-229-7
7	8266	93.0	1715	14	US-10-021-571-2
8	8266	93.0	1715	14	US-10-117-229-6
9	5899.5	66.4	1184	14	US-10-117-229-3
10	5899.5	66.4	1184	14	US-10-117-229-9
11	3050.5	34.3	705	14	US-10-106-698-6378
12	2851.5	32.1	551	9	US-09-835-788A-17
13	2851.5	32.1	551	14	US-10-175-042-17

14	2638	29.7	513	14	US-10-149-819-9	Sequence 9, Appli
15	2379.5	26.8	1498	14	US-10-021-571-8	Sequence 8, Appli
16	1832.5	20.6	1398	14	US-10-021-571-6	Sequence 6, Appli
17	553	6.2	109	10	US-09-986-480-371	Sequence 371, App
18	543.5	6.1	657	15	US-10-104-047-2529	Sequence 2529, Ap
19	519	5.8	1724	9	US-09-984-899-43	Sequence 43, Appli
20	506	5.7	1330	15	US-10-108-260A-3237	Sequence 3237, Ap
21	474	5.3	1762	14	US-10-205-194-117	Sequence 117, App
22	466.5	5.3	1188	16	US-10-408-765A-1311	Sequence 1311, Ap
23	448	5.0	747	15	US-10-094-749-1924	Sequence 1924, Ap
24	444.5	5.0	919	15	US-10-108-260A-4122	Sequence 4122, Ap
25	443.5	5.0	3913	15	US-10-334-143-45	Sequence 45, Appli
26	438.5	4.9	1053	15	US-10-291-172-343	Sequence 343, App
27	438.5	4.9	1053	15	US-10-221-278-343	Sequence 343, App
28	437	4.9	740	9	US-09-835-788A-12	Sequence 12, Appli
29	437	4.9	740	14	US-10-175-042-12	Sequence 12, Appli
30	437	4.9	1569	15	US-10-275-595A-31	Sequence 31, Appli
31	409.5	4.6	1431	15	US-10-352-684A-38	Sequence 38, Appli
32	409.5	4.6	1431	15	US-10-045-400C-5	Sequence 5, Appli
33	409.5	4.6	1431	15	US-10-295-027-1235	Sequence 1235, Ap
34	409	4.6	367	15	US-10-250-613-6	Sequence 6, Appli
35	408.5	4.6	426	9	US-09-908-711-70	Sequence 70, Appli
36	395.5	4.5	1360	16	US-10-473-574-12	Sequence 12, Appli
37	390	4.4	1554	14	US-10-301-822-93	Sequence 93, Appli
38	390	4.4	1697	15	US-10-112-944-370	Sequence 370, App
39	388	4.4	622	15	US-10-112-944-345	Sequence 345, App
40	385.5	4.3	765	14	US-10-128-174-3	Sequence 3, Appli
41	385.5	4.3	765	14	US-10-128-174-34	Sequence 34, Appli
42	385.5	4.3	765	14	US-10-128-174-35	Sequence 35, Appli
43	385.5	4.3	765	14	US-10-128-174-36	Sequence 36, Appli
44	385.5	4.3	765	14	US-10-128-174-37	Sequence 37, Appli
45	385.5	4.3	765	14	US-10-128-174-38	Sequence 38, Appli

ALIGNMENTS

RESULT 1  
US-10-021-571-4  
; Sequence 4, Application US/10021571  
; Publication No. US20030166056A1  
; GENERAL INFORMATION:  
; APPLICANT: CHAO, Moses V.  
; APPLICANT: KONG, Haeyoung  
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN  
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTI  
; FILE REFERENCE: CHAO11A  
; CURRENT APPLICATION NUMBER: US/10/021.571  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/256,909  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1715  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-021-571-4

Query Match	100.0%;	Score 8884;	DB 14;	Length 1715;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1715;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERNEGQTPMTAAEQGNLEIVKELIKN	60	
Db	1	MSVLISQSVINYVEENIPALKALLEKCKVDNERNEGQTPMTAAEQGNLEIVKELIKN	60	
Qy	61	GANCLNLDLNTALISASKEGHVHVEELLKCGVNLNHRDMGWTALMWAQYKGRDVV	120	
Db	61	GANCLNLDLNTALISASKEGHVHVEELLKCGVNLNHRDMGWTALMWAQYKGRDVV	120	
Qy	121	ELLISHGANFSVTGLQYSVPYIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLTWWAAR	180	
Db	121	ELLISHGANFSVTGLQYSVPYIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLTWWAAR	180	

121 ELLLSHGANPSVTGLQSVVPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAAR 180  
181 KGHLEVCVKKHLLANGADVDOEGANSM TALIVAVKGGYTQSVEILKENPNVNLTDKGNTA 240  
181 KGHLEVCVKKHLLANGADVDOEGANSM TALIVAVKGGYTQSVEILKENPNVNLTDKGNTA 240  
241 LMIASKEGHEITEIVQDILLDAGTYNNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
241 LMIASKEGHEITEIVQDILLDAGTYNNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
301 QDNKTALYNAVEKGNATWVDILQCNPDTEICTKGETPLIKATKMRNIEVVELLDKGA 360  
301 QDNKTALYNAVEKGNATWVDILQCNPDTEICTKGETPLIKATKMRNIEVVELLDKGA 360  
361 KVSADVKKGDTPLHIAIRGSRKLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420  
361 KVSADVKKGDTPLHIAIRGSRKLAELLRNPKDGRLLYRPNKAGETPNYIDCSHOKSIL 420  
421 TQIFGARHLSPTEIDGMDLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKGKSFLLKKL 480  
421 TQIFGARHLSPTEIDGMDLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKGKSFLLKKL 480  
481 EDEMKTFAGQIEPLPQFQSWLIVFLTLCCGGIGLLFAFTVHPNGLGIAVLSFLALLYIF 540  
481 EDEMKTFAGQIEPLPQFQSWLIVFLTLCCGGIGLLFAFTVHPNGLGIAVLSFLALLYIF 540  
541 FIVYIFGRRREGESWNAWVLSRLARHIGYLELLKLMFVNPPPELPEOTTKALPVRPLF 600  
541 FIVYIFGRRREGESWNAWVLSRLARHIGYLELLKLMFVNPPPELPEOTTKALPVRPLF 600  
601 TDYNRLLSVGGETSLEAMTATLSDACEREFGLATRLFRVFKTEDTQGGKKWKTCCLPS 660  
601 TDYNRLLSVGGETSLEAMTATLSDACEREFGLATRLFRVFKTEDTQGGKKWKTCCLPS 660  
661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGGLAFVLCNRTWQVLDLSLL 720  
661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGGLAFVLCNRTWQVLDLSLL 720  
721 NSORKELHNAASKHLKSEGFMMVKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDAC 780  
721 NSORKELHNAASKHLKSEGFMMVKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDAC 780  
781 EODKVLQMLDTRVFLSKGPFPTAIPASDPHIIKAINQNLNSVLRDSNNGHDYMNIVH 840  
781 EODKVLQMLDTRVFLSKGPFPTAIPASDPHIIKAINQNLNSVLRDSNNGHDYMNIVH 840  
841 LPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGIGQEDADRRVSQNSLGEWTKLGSKTALN 900  
841 LPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGIGQEDADRRVSQNSLGEWTKLGSKTALN 900  
901 RRDYTRRQWORTITRQMSFDLTKLVTEWDFSDISPOWRRLLNIVSVTGRLLRANQIS 960  
901 RRDYTRRQWORTITRQMSFDLTKLVTEWDFSDISPOWRRLLNIVSVTGRLLRANQIS 960  
961 FNDWRLASINLTQOPPYRTSMILLYLETEGIPDQWTLKTIYERISKNIPTTKQVEPLL 1020  
961 FNDWRLASINLTQOPPYRTSMILLYLETEGIPDQWTLKTIYERISKNIPTTKQVEPLL 1020  
1021 EIDGDIRNFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGLAY 1080  
1021 EIDGDIRNFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGLAY 1080  
1081 PPLPLHGGPPRAPSGVSPVSCSSTS FNGPAGGVSPQPHSSYSGMTGQPHPPYNRG 1140  
1081 PPLPLHGGPPRAPSGVSPVSCSSTS FNGPAGGVSPQPHSSYSGMTGQPHPPYNRG 1140  
1141 SGAPGPVLLNSLVNDVACEKLKQIEGLDOSMLPOYCTTIKKANINGVLAQCNDLEK 1200  
1141 SGAPGPVLLNSLVNDVACEKLKQIEGLDOSMLPOYCTTIKKANINGVLAQCNDLEK 1200  
1201 KEMNNMFGDWHIFRSTVLEMRNAESHVWPEDRFLSESSSGPAPGEPARRASHNELPHT 1260  
1201 KEMNNMFGDWHIFRSTVLEMRNAESHVWPEDRFLSESSSGPAPGEPARRASHNELPHT 1260

1261 ELSSTQPTYLNFSEBELNTLGLDEGAPRHSNLSWOSOTRRTPSLSLNSODSSIEISKLT 1320  
1261 ELSSTQPTYLNFSEBELNTLGLDEGAPRHSNLSWOSOTRRTPSLSLNSODSSIEISKLT 1320  
1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGSIIHNSLQEK 1380  
1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGSIIHNSLQEK 1380  
1381 KQSEPKDDGRKSFMLKRGDVIDYSSGVSTVNDASPLDPIITBEDEKSDQSGKLLPGKKS 1440  
1381 KQSEPKDDGRKSFMLKRGDVIDYSSGVSTVNDASPLDPIITBEDEKSDQSGKLLPGKKS 1440  
1441 SRSISLFOYDLKLGSLRYQKLPSEDESGTSEEDNTPLKKDDKDKAEKVVERPKSP 1500  
1441 SRSISLFOYDLKLGSLRYQKLPSEDESGTSEEDNTPLKKDDKDKAEKVVERPKSP 1500  
1501 EHSASPIRTFIIKAXEYLSDALDDKDDSDSGVRSSESSPNHSLHNEVADDSQLEKANLIE 1560  
1501 EHSASPIRTFIIKAXEYLSDALDDKDDSDSGVRSSESSPNHSLHNEVADDSQLEKANLIE 1560  
1561 LEDDSHGKRGIPHSLSGLQDPIIARMSICSDKXKSPSECSLIASSPENWPAQKAYNL 1620  
1561 LEDDSHGKRGIPHSLSGLQDPIIARMSICSDKXKSPSECSLIASSPENWPAQKAYNL 1620  
1621 NRTPTVTTLNNSAPANRANQNFDEMEGIRETSOVLIRPSSSPNPPTTIONENLKSMTKR 1680  
1621 NRTPTVTTLNNSAPANRANQNFDEMEGIRETSOVLIRPSSSPNPPTTIONENLKSMTKR 1680  
1681 SORSSYTRLSKDPPELHAAASSESTGFGEERESIL 1715  
1681 SORSSYTRLSKDPPELHAAASSESTGFGEERESIL 1715

## RESULT 2

US-10-117-229-2  
; Sequence 2, Application US/10117229  
; Publication No. US20030190625A1  
; GENERAL INFORMATION:  
; APPLICANT: OriGene Technologies, Inc.  
; TITLE OF INVENTION: Human Kidins220Pc  
; FILE REFERENCE: 9U 104 R1  
; CURRENT APPLICATION NUMBER: US/10/117,229  
; CURRENT FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 2  
; LENGTH: 1715  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-117-229-2

Query Match 100.0%; Score 8884; DB 14; Length 1715;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSVLISQSIVNYVEENIPALKALLEKCKVDVERNECGQTPLMIAAQNLEIVKELIKN 60  
1 MSVLISQSIVNYVEENIPALKALLEKCKVDVERNECGQTPLMIAAQNLEIVKELIKN 60  
61 GANCNLEDLNNWTALISASKEGHVHIVBELKCGVNLHRDGGWTALMWACVKGRTDVV 120  
61 GANCNLEDLNNWTALISASKEGHVHIVBELKCGVNLHRDGGWTALMWACVKGRTDVV 120  
121 ELLLSHGANPSVTGLQSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAAR 180  
121 ELLLSHGANPSVTGLQSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAAR 180  
181 KGHLEVCVKKHLLANGADVDOEGANSM TALIVAVKGGYTQSVEILKENPNVNLTDKGNTA 240  
181 KGHLEVCVKKHLLANGADVDOEGANSM TALIVAVKGGYTQSVEILKENPNVNLTDKGNTA 240  
241 LMIASKEGHEITEIVQDILLDAGTYNNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300



Db 241 LMIASKEGHEIVQDLIDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300  
Qy 301 QDNKTALYAWVEKGNATVRDILQCNPDTEICTKDGTEPLIKATQWNIIEVVELLDKGA 360  
Db 301 QDNKTALYAWVEKGNATVRDILQCNPDTEICTKDGTEPLIKATQWNIIEVVELLDKGA 360  
Qy 361 KVSADVKKGGDTPHIAIRGRSKLAELLLRNPKDGRLLYRPNKAGETPNYDCSHQKSL 420  
Db 361 KVSADVKKGGDTPHIAIRGRSKLAELLLRNPKDGRLLYRPNKAGETPNYDCSHQKSL 420  
Qy 421 TOIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAGWGSKSFLKKL 480  
Db 421 TOIFGARHLSPTETDGMGLYDLYSSALADILSEPTMQPPICVGLYAGWGSKSFLKKL 480  
Qy 481 EDEMKTFAQOIIEPLFQFQSWLIVFLTLCCGLGLLPAFTVHPNLGIAVLSFLALLYIF 540  
Db 481 EDEMKTFAQOIIEPLFQFQSWLIVFLTLCCGLGLLPAFTVHPNLGIAVLSFLALLYIF 540  
Qy 541 FIVYFGRREGESNNWAVLSTRARHIGYLELLKLMFVNPPELPEQTTKALPVRLF 600  
Db 541 FIVYFGRREGESNNWAVLSTRARHIGYLELLKLMFVNPPELPEQTTKALPVRLF 600  
Qy 601 TDYRLSVGGETSLAEMIATLSACEREFGLATRLPRVFKTEDTQKKKWKTCCLPS 660  
Db 601 TDYRLSVGGETSLAEMIATLSACEREFGLATRLPRVFKTEDTQKKKWKTCCLPS 660  
Qy 661 FVIFLFIIGCIISGIIITLAIIRVDPKHLTVNAVLISIASVVGLAFVLCRTWQVLDL 720  
Db 661 FVIFLFIIGCIISGIIITLAIIRVDPKHLTVNAVLISIASVVGLAFVLCRTWQVLDL 720  
Qy 721 NSQRKRLHNAASKLHLKSEGFMVKYKCEVELMARMAKTIDSFQNTQRLVVIIDGLDAC 780  
Db 721 NSQRKRLHNAASKLHLKSEGFMVKYKCEVELMARMAKTIDSFQNTQRLVVIIDGLDAC 780  
Qy 781 BQDKVQLMDTVRVLFSKGPPIAIPASDPHIIIIKAINQNLNLSVRDQSNINGHDYMRNIVH 840  
Db 781 BQDKVQLMDTVRVLFSKGPPIAIPASDPHIIIIKAINQNLNLSVRDQSNINGHDYMRNIVH 840  
Qy 841 LPVFLNLSRGLSNARKFLVTSATNGDVPKSDTGTQEDADRVSONSGEMTKLASKTALN 900  
Db 841 LPVFLNLSRGLSNARKFLVTSATNGDVPKSDTGTQEDADRVSONSGEMTKLASKTALN 900  
Qy 901 RRDYTRRQMQRTITROMSFDLTKLLVTEDFWSDISPTQMERLNLNLSVTRGLLRANQIS 960  
Db 901 RRDYTRRQMQRTITROMSFDLTKLLVTEDFWSDISPTQMERLNLNLSVTRGLLRANQIS 960  
Qy 961 FNWDLASWINLTQWQPYRTSWLILYLEETEGIPDQMTLKIYIRISKNIPTTKDVEPLL 1020  
Db 961 FNWDLASWINLTQWQPYRTSWLILYLEETEGIPDQMTLKIYIRISKNIPTTKDVEPLL 1020  
Qy 1021 EIDGDIRNFVFLASRTPVLAARDVKVPLCTVNLDPKREIIADVRAAREQISIGGLAY 1080  
Db 1021 EIDGDIRNFVFLASRTPVLAARDVKVPLCTVNLDPKREIIADVRAAREQISIGGLAY 1080  
Qy 1081 PPLPLHEGPPAPAGYSGOPSPVCSSTSNFGAGVSPQPHSSYSGMTGPQPFYNRG 1140  
Db 1081 PPLPLHEGPPAPAGYSGOPSPVCSSTSNFGAGVSPQPHSSYSGMTGPQPFYNRG 1140  
Qy 1141 SGAPAGPVVLNLSNDAVCEKLOIEGLDQSMPLPOYCTTIKKANINGRVLACQNDLX 1200  
Db 1141 SGAPAGPVVLNLSNDAVCEKLOIEGLDQSMPLPOYCTTIKKANINGRVLACQNDLX 1200  
Qy 1201 KEMMNFGDWHLFRSTVLEMRNASHVVPDPRLSESSSGPAPGHPARRASHNELPHT 1260  
Db 1201 KEMMNFGDWHLFRSTVLEMRNASHVVPDPRLSESSSGPAPGHPARRASHNELPHT 1260  
Qy 1261 ELSSTQPTVLNFSFEELNTLGLDGAHNSLWSQOTRTPPSLSSNSQSSSIEISKLT 1320  
Db 1261 ELSSTQPTVLNFSFEELNTLGLDGAHNSLWSQOTRTPPSLSSNSQSSSIEISKLT 1320  
Qy 1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYTMGSSSGGSIHNSLEQK 1380

Db 1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYTMGSSSGGSIHNSLEQK 1380  
Qy 1381 KDSPEKPDGGRKSFIMKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKS 1440  
Db 1381 KDSPEKPDGGRKSFIMKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKS 1440  
Qy 1441 SERSSLFQTDLKLKSGSLRYOKLPDSEDESGTESDNTPLLKDDKDKRKAEGKVERVPKSP 1500  
Db 1441 SERSSLFQTDLKLKSGSLRYOKLPDSEDESGTESDNTPLLKDDKDKRKAEGKVERVPKSP 1500  
Qy 1501 EHSAPERTFTKAKEXYLSALLDKKSDSDSGVRSESSPNHSLHNEVADDSQLEKANLIE 1560  
Db 1501 EHSAPERTFTKAKEXYLSALLDKKSDSDSGVRSESSPNHSLHNEVADDSQLEKANLIE 1560  
Qy 1561 LEDDSHSGKRGIPHSLSGLQDPIIARMSICSEDKSPSECSLIASSPENNWPACOKAYNL 1620  
Db 1561 LEDDSHSGKRGIPHSLSGLQDPIIARMSICSEDKSPSECSLIASSPENNWPACOKAYNL 1620  
Qy 1621 NRTPTVTNLNNSAPANRANQNFDEMGIRETSQVILRPSSSPNPTTIONENLKSMTHKR 1680  
Db 1621 NRTPTVTNLNNSAPANRANQNFDEMGIRETSQVILRPSSSPNPTTIONENLKSMTHKR 1680  
Qy 1681 SORSSTYRLSKDPPHLLHAAASSESTGFGEERESIL 1715  
Db 1681 SORSSTYRLSKDPPHLLHAAASSESTGFGEERESIL 1715

## RESULT 3

US-10-473-574-26  
; Sequence 26, Application US/10473574  
; Publication No. US20040116670A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION; HAPALIA, April J.A.;  
; APPLICANT: TANG, Y. Tom; YUE, Henry;  
; APPLICANT: KHAN, Farrah A.; ISON, Craig H.;  
; APPLICANT: BAUGHN, Mariah R.; WARREN, Bridget A.;  
; APPLICANT: DUGGAN, Brendan M.; THANGAVELU, Kavitha;  
; APPLICANT: HONCHELL, Cynthia D.; AZIMZAI, Valda;  
; APPLICANT: ELLIOTT, Vicki S.; BURFORD, Neil;  
; APPLICANT: DING, Li; YUE, HuiBin;  
; APPLICANT: BECHA, Shanya; EMERLING, Brooke M.;  
; APPLICANT: RICHARDSON, Thomas W.; LEE, Soo Yeun;  
; APPLICANT: BANDMAN, Olga; LAL, Preeti G.;  
; APPLICANT: LEE, Sally; GIETZEN, Kimberly J.;  
; APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;  
; APPLICANT: LEE, Ernestine A.; SWARNAKAR, Anita;  
; APPLICANT: KING, HuiJun Z.; JONES, Karen Anne  
; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0918 USN  
; CURRENT APPLICATION NUMBER: US/10/473,574  
; CURRENT FILING DATE: 2003-09-29  
; PRIOR APPLICATION NUMBER: PCT/US02/09288  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER: US 60/294,451  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: US 60/291,870  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/290,518  
; PRIOR FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: US 60/288,609  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/283,769  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/281,323  
; PRIOR FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/280,508  
; PRIOR FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PERL Program  
; SEQ ID NO 26  
; LENGTH: 1715  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; FEATURE: ; NAME/KEY: misc feature ; OTHER INFORMATION: Incyte ID No: 2755454CDI US-10-473-574-26									
Query Match 100.0%; Score 8884; DB 16; Length 1715; Best Local Similarity 100.0%; Pred. No. 0; Matches 1715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MSVLSQSUNVYVEENIPALKALLSKCDVDNRNCGGOTPLMAAEQGNLEIVKELIKN	60						
DB	1	MSVLSQSUNVYVEENIPALKALLSKCDVDNRNCGGOTPLMAAEQGNLEIVKELIKN	60						
QY	61	GANCNLEDLNTWALISASKEGHVHIVEELLKGVNLEHRDMGWTALWACVKGRTDVV	120						
DB	61	GANCNLEDLNTWALISASKEGHVHIVEELLKGVNLEHRDMGWTALWACVKGRTDVV	120						
QY	121	ELLSSHGANSVTVGLQVSVYPIIWAAGRGHADIHLLQLNGAKVNCSDKYGTTPLVWAAR	180						
DB	121	ELLSSHGANSVTVGLQVSVYPIIWAAGRGHADIHLLQLNGAKVNCSDKYGTTPLVWAAR	180						
QY	181	KHLECVKHLANGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKQNTA	240						
DB	181	KHLECVKHLANGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKQNTA	240						
QY	241	LMTASKEGTEIYODLIDAGTYNNIPDRSGDTVLI GAVRGHVEIVRALLQKYADIDIRG	300						
DB	241	LMTASKEGTEIYODLIDAGTYNNIPDRSGDTVLI GAVRGHVEIVRALLQKYADIDIRG	300						
QY	301	QDNKTALYWAVEKGNATWVDIILQCPNDEICTKDGTEPLIKATKRNIEVVELLDKGA	360						
DB	301	QDNKTALYWAVEKGNATWVDIILQCPNDEICTKDGTEPLIKATKRNIEVVELLDKGA	360						
QY	361	KVSAVDKGDTPHIAIRGSRKLABELLRNPKDGRLLYRPNKAGETPNYINDCSHOKSIL	420						
DB	361	KVSAVDKGDTPHIAIRGSRKLABELLRNPKDGRLLYRPNKAGETPNYINDCSHOKSIL	420						
QY	421	TOIFGARHLSPTTGDGMGLYDLYSSNALADILSEPTMOPPICVGLYAQWGSFGSLKLL	480						
DB	421	TOIFGARHLSPTTGDGMGLYDLYSSNALADILSEPTMOPPICVGLYAQWGSFGSLKLL	480						
QY	481	EDEMKTFAGQOIEPLFQFQSWLIVFLTLILCGGLLFAFTVHPNLGIAVSLSLALLIYF	540						
DB	481	EDEMKTFAGQOIEPLFQFQSWLIVFLTLILCGGLLFAFTVHPNLGIAVSLSLALLIYF	540						
QY	541	FIVIFGGRREGESWNAWYLSRLARHIGYLELLKLMFVNPPELPEOTTKALPVRFEL	600						
DB	541	FIVIFGGRREGESWNAWYLSRLARHIGYLELLKLMFVNPPELPEOTTKALPVRFEL	600						
QY	601	TDYRLSSVGGETSLEAMIAITLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCLPS	660						
DB	601	TDYRLSSVGGETSLEAMIAITLSDACEREFGLATRLFRVFKTEDTQKKKKWKTCLPS	660						
QY	661	FVIFLFIIGCIIGITLLALFRVDPKHLTVNAVLIASVVGGLAVLNCRTWQVLDLSLL	720						
DB	661	FVIFLFIIGCIIGITLLALFRVDPKHLTVNAVLIASVVGGLAVLNCRTWQVLDLSLL	720						
QY	721	NSQKRLHNAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQLRVIIIDGLDAC	780						
DB	721	NSQKRLHNAASKLHLKSEGFMKVLKCEVELMARMAKTIDSFTQNTQLRVIIIDGLDAC	780						
QY	781	EQDKVQLMDTVTVLPSKGFPIAFASDPHIIKAINONLSVLRDSNINCHDYMNRIVH	840						
DB	781	EQDKVQLMDTVTVLPSKGFPIAFASDPHIIKAINONLSVLRDSNINCHDYMNRIVH	840						
QY	841	LPVFLNSRGLSNARKFLVTSATNGDVPCTDTGQEDADRRVSQNSLIGEMTKLGSKTALN	900						
DB	841	LPVFLNSRGLSNARKFLVTSATNGDVPCTDTGQEDADRRVSQNSLIGEMTKLGSKTALN	900						
QY	901	RRDTYRRRQWRTITRQMSFDLTKLLVTDWDFSDISPTQWRRLLNIVSVTGRLLRANQIS	960						
DB	901	RRDTYRRRQWRTITRQMSFDLTKLLVTDWDFSDISPTQWRRLLNIVSVTGRLLRANQIS	960						

RESULT 4  
 US-10-117-229-4  
 ; Sequence 4, Application US/10117229  
 ; Publication No. US20030190625A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Origene Technologies, Inc.  
 ; TITLE OF INVENTION: Human Kidins220Pc  
 ; FILE REFERENCE: 9U 104 R1  
 ; CURRENT APPLICATION NUMBER: US/10/117,229  
 ; CURRENT FILING DATE: 2002-04-08  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 1771  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-117-229-4  
 Query Match 99.4%; Score 8830; DB 14; Length 1771;  
 Best Local Similarity 96.7%; Pred. No. 0;

Matches 1714; Conservative 0; Mismatches 0; Indels 58; Gaps 2;			
Qy	1	MSVLISQSVINYVEENIPALKALLECKVDNERNECQOTPLMTAAEQNLEIVKELIKN	60
Db	1	MSVLISQSVINYVEENIPALKALLECKVDNERNECQOTPLMTAAEQNLEIVKELIKN	60
Qy	61	GANCNLELDNWTALISAKSGHVHIVBELLKCGVNLEHRMGWWTALMWACYKGRDGV	120
Db	61	GANCNLELDNWTALISAKSGHVHIVBELLKCGVNLEHRMGWWTALMWACYKGRDGV	120
Qy	121	ELLASHGANPSVTGLOYSVYPIIWAAGRGHADI VHLILONGAKVNCSDKYGTTPLVWAAR	180
Db	121	ELLASHGANPSVTGL-YSVYPIIWAAGRGHADI VHLILONGAKVNCSDKYGTTPLVWAAR	179
Qy	181	KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTQSVKEILKRNPNVNLTKDGNTA	240
Db	180	KGHLECVKHLAMGADVDOEGANSMTALIVAKGYTQSVKEILKRNPNVNLTKDGNTA	239
Qy	241	LMIASKEGHTI VODLLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG	300
Db	240	LMIASKEGHTI VODLLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG	299
Qy	301	QDNKTALYAVEKGNATVRILOCPNPTETCTKDGTEPLIKATKRNIEVVELLLDKGA	360
Db	300	QDNKTALYAVEKGNATVRILOCPNPTETCTKDGTEPLIKATKRNIEVVELLLDKGA	359
Qy	361	KVSAVDKGDTPHIAIRGRSKLAELLRNPKDGLLYRNPKAGETPNYDCSHQSKIL	420
Db	360	KVSAVDKGDTPHIAIRGRSKLAELLRNPKDGLLYRNPKAGETPNYDCSHQSKIL	419
Qy	421	TOIFGARHLSPTETDGMGLDYSSALADILSEPTMQPPICVGLYAWGSGKSFLLKKL	480
Db	420	TOIFGARHLSPTETDGMGLDYSSALADILSEPTMQPPICVGLYAWGSGKSFLLKKL	479
Qy	481	EDEMKTFAGQOI EPLFQPSWLI VFLTL LCGGLLFAFTVHPNLGIAVLSFLALIYIF	540
Db	480	EDEMKTFAGQOI EPLFQPSWLI VFLTL LCGGLLFAFTVHPNLGIAVLSFLALIYIF	539
Qy	541	FIVYFGRRGESNNWAWLSTRARHIGYLELLKL MFVNPPELPBOTTALPVRFPLF	600
Db	540	FIVYFGRRGESNNWAWLSTRARHIGYLELLKL MFVNPPELPBOTTALPVRFPLF	599
Qy	601	TDYNRLSVGGTSLAEMIATLSACEREFGFLATRLFRVFKETDTGKKWKKTCCPLPS	660
Db	600	TDYNRLSVGGTSLAEMIATLSACEREFGFLATRLFRVFKETDTGKKWKKTCCPLPS	659
Qy	661	FVIFLFIIGCIISGITLLAIFRVPDPKHLTVNAVLI SIASVVGFLAPVLCRTWQVLDL	720
Db	660	FVIFLFIIGCIISGITLLAIFRVPDPKHLTVNAVLI SIASVVGFLAPVLCRTWQVLDL	719
Qy	721	NSQKRLHNAASKHLKSGFMVLCVVELMARMAKTIDSPQNTQRLVVIIDGLDAC	780
Db	720	NSQKRLHNAASKHLKSGFMVLCVVELMARMAKTIDSPQNTQRLVVIIDGLDAC	779
Qy	781	BQDKVQLMDTVRVLFSGPPIAIFASDPHII IKAIONLNSVLDRDSNINGHDYRNIVH	840
Db	780	BQDKVQLMDTVRVLFSGPPIAIFASDPHII IKAIONLNSVLDRDSNINGHDYRNIVH	839
Qy	841	LPVFLNSRGLSNARKFLVTSATNGDVPDCTGTQEDADRVSQNSLGEMTKLGSKTALN	900
Db	840	LPVFLNSRGLSNARKFLVTSATNGDVPDCTGTQEDADRVSQNSLGEMTKLGSKTALN	899
Qy	901	RRDYRRQORTITROMSPDLTKLVTEDFWSDISPTQMERLLNIVSVTGRLLRANQIS	960
Db	900	RRDYRRQORTITROMSPDLTKLVTEDFWSDISPTQMERLLNIVSVTGRLLRANQIS	959
Qy	961	FNWDLASWINLTEQWPRTSWLILYLETREGIPDQMTLKIYERISKNIPTTKDVEPLL	1020
Db	960	FNWDLASWINLTEQWPRTSWLILYLETREGIPDQMTLKIYERISKNIPTTKDVEPLL	1019
Qy	1021	EIDGDIRNFEVFLSSRTPVLDVQKFLPCTVNLDPKRLREIIADVRAAREQISIGGLAY	1080
Db	1020	EIDGDIRNFEVFLSSRTPVLDVQKFLPCTVNLDPKRLREIIADVRAAREQISIGGLAY	1079

Qy	1081	PPLPLHREPPRAPSGYQPPSCVSTSNFGPAGVSPQPHSSYYSQMTGPQHPFYN--	1138
Db	1080	PPLPLHREPPRAPSGYQPPSCVSTSNFGPAGVSPQPHSSYYSQMTGPQHPFYNRP	1139
Qy	1139	-----RSGSP	1143
Db	1140	FFAPLYTPRYPGSQHLSIRPSVKTSLPRDQNNGLEVIKEDAAEGLSSPTDSSRGSGP	1199
Qy	1144	APGPVLLNSLVNDAVCEKLOIEGLDQSMLOPYCTTIKKANINGRVLAQCNIDELKKEM	1203
Db	1200	APGPVLLNSLVNDAVCEKLOIEGLDQSMLOPYCTTIKKANINGRVLAQCNIDELKKEM	1259
Qy	1204	NNNFGDWHLFSTVLEMRNAESHVVPDRFLSESSSGPAPHGPPARRASHNELPHTELS	1263
Db	1260	NNNFGDWHLFSTVLEMRNAESHVVPDRFLSESSSGPAPHGPPARRASHNELPHTELS	1319
Qy	1264	SQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSOTRTPPSLSSNQSSSIISKLTQKV	1323
Db	1320	SQTPYTLNFSPEELNTLGLDEGAPRHSNLSWQSOTRTPPSLSSNQSSSIISKLTQKV	1379
Qy	1324	QAEYRDAYREVI AOWSOL EGGPGSTTISGRSSPHSTYYMGOSSGGSIHSLNLEQKGD	1383
Db	1380	QAEYRDAYREVI AOWSOL EGGPGSTTISGRSSPHSTYYMGOSSGGSIHSLNLEQKGD	1439
Qy	1384	EPKPDGGRKSPMKRGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKSSE	1443
Db	1440	EPKPDGGRKSPMKRGDVIDYSSSGVSTNDASPLDPITEEDEKSDQSGSKLLPGKKSSE	1499
Qy	1444	SSLFQDTLKLKGGSLRYQKLPSEDESGTESDNTPLLKDDKRAEKGVKVERVPKSPH	1503
Db	1500	SSLFQDTLKLKGGSLRYQKLPSEDESGTESDNTPLLKDDKRAEKGVKVERVPKSPH	1559
Qy	1504	AEPIRTFKAYEVLSDALLDKKSSDSGVRSSESPNHSNLHNEVADDLSOLEKANLIE	1563
Db	1560	AEPIRTFKAYEVLSDALLDKKSSDSGVRSSESPNHSNLHNEVADDLSOLEKANLIE	1619
Qy	1564	DSHSGKRGIPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNT	1623
Db	1620	DSHSGKRGIPHSLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACQKAYNLNT	1679
Qy	1624	PSTVTLNNSGAPANRANQNFDEMEGIRETSQVILRPSPPNPTTIQNLKSMTHKRSQR	1683
Db	1680	PSTVTLNNSGAPANRANQNFDEMEGIRETSQVILRPSPPNPTTIQNLKSMTHKRSQR	1739
Qy	1684	SSYTRLKDPPELHAAASSESTGFEERESIL	1715
Db	1740	SSYTRLKDPPELHAAASSESTGFEERESIL	1771

RESULT 5

US-10-117-229-11  
; Sequence 11, Application US/10117229  
; Publication No. US20030190625A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc.  
; TITLE OF INVENTION: Human Kidine220pc  
; FILE REFERENCE: 9U 104 R1  
; CURRENT APPLICATION NUMBER: US/10/117,229  
; CURRENT FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 1771  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-117-229-11

Query Match 99.4%; Score 8830; DB 14; Length 1771;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 1714; Conservative 0; Mismatches 0; Indels 58; Gaps 2;

Qy 1 MSVLISQSVINYVEENIPALKALLECKVDNERNECQOTPLMTAAEQNLEIVKELIKN 60

Db 1 MSVLISQSVINYVEENIPALKALLECKDVDERNECCQTPLMAAEOGNLEIVKELIKN 60  
QY 61 GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHRDMGWTALMAWACYKGRDVV 120  
Db 61 GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHRDMGWTALMAWACYKGRDVV 120  
QY 121 ELLLSHGANDSVTGLQSVYVPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPVWAAR 180  
Db 121 ELLLSHGANDSVTGLQSVYVPIIWAAGRGHADIVHLLQNGAKVNCSDKYGTTPVWAAR 179  
QY 181 KGHLECVKHLANGADVDOGANSMTALIVAVKGYTQSVKEILKKNPNVNLTKDGN 240  
Db 180 KGHLECVKHLANGADVDOGANSMTALIVAVKGYTQSVKEILKKNPNVNLTKDGN 239  
QY 241 LMIASKEGHTIEIVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300  
Db 240 LMIASKEGHTIEIVQDLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299  
QY 301 QDNKTALYWAVERKGNATMVRDILQCNPDTEICTKDGTEPLIKATKWRNIEVVELLDKGA 360  
Db 300 QDNKTALYWAVERKGNATMVRDILQCNPDTEICTKDGTEPLIKATKWRNIEVVELLDKGA 359  
QY 361 KVSADVKKGDTPLHIAIRGSRKLAELLLRNPXGRLLYRPNKAGETPNYIDCSHOKSIL 420  
Db 360 KVSADVKKGDTPLHIAIRGSRKLAELLLRNPXGRLLYRPNKAGETPNYIDCSHOKSIL 419  
QY 421 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFGSLKKL 480  
Db 420 TQIFGARHLSPTETDGMGLYDLYSSALADILSEPTWQPPICVGLYAQWGSFGSLKKL 479  
QY 481 EDEMKTIFAGQIIEPLFQFQSWLIVFLTLICGGLGLLFAFTVHNLGIAVSLFALLIYIF 540  
Db 480 EDEMKTIFAGQIIEPLFQFQSWLIVFLTLICGGLGLLFAFTVHNLGIAVSLFALLIYIF 539  
QY 541 FIVYFGRREGSNNWVLSRLARHIGYLELLKLMFVNPELPEQTTKALPVRELF 600  
Db 540 FIVYFGRREGSNNWVLSRLARHIGYLELLKLMFVNPELPEQTTKALPVRELF 599  
QY 601 TDYRNLSSVGGETSLEAMITATLSDACEREFGLATRLFRVFKTEDTQGGKKWKTKCLPS 660  
Db 600 TDYRNLSSVGGETSLEAMITATLSDACEREFGLATRLFRVFKTEDTQGGKKWKTKCLPS 659  
QY 661 FVIFLFTIGCIIISGITLLAIFRVDPKHLTNAVLISIASVVGGLAFVNLCKRTWQVLDL 720  
Db 660 FVIFLFTIGCIIISGITLLAIFRVDPKHLTNAVLISIASVVGGLAFVNLCKRTWQVLDL 719  
QY 721 NSQKRRLHNAASKLHKLKSGFWMKVKCEVELMARWAKTIDSTFQNTQRLVLIIDGLDAC 780  
Db 720 NSQKRRLHNAASKLHKLKSGFWMKVKCEVELMARWAKTIDSTFQNTQRLVLIIDGLDAC 779  
QY 781 EQDKVLQMLDTRVFLSKGFFIAIFASDPHIIKAINQNLSVLRDSNNGHDMYRNIVH 840  
Db 780 EQDKVLQMLDTRVFLSKGFFIAIFASDPHIIKAINQNLSVLRDSNNGHDMYRNIVH 839  
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDTTIGQEDADRVSQNSIGEMTKLGSKTALN 900  
Db 840 LPVFLNSRGLSNARKFLVTSATNGDVPSCDTTIGQEDADRVSQNSIGEMTKLGSKTALN 899  
QY 901 RRDYRRRQORTITRQMSFDLTKLVTEWDFSDISPTQMRRLNTVSVTGRLLRANQIS 960  
Db 900 RRDYRRRQORTITRQMSFDLTKLVTEWDFSDISPTQMRRLNTVSVTGRLLRANQIS 959  
QY 961 FNNWDLASWINLQEQPYRTSWLILYLETEGIPDQMTLXTIYERISKNIPTTKDVEPLL 1020  
Db 960 FNNWDLASWINLQEQPYRTSWLILYLETEGIPDQMTLXTIYERISKNIPTTKDVEPLL 1019  
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKREIITADVRAAREQISIGLAY 1080  
Db 1020 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKREIITADVRAAREQISIGLAY 1079  
QY 1081 PPLPLHEGPPRAPSGYSQPSVCSSTSFGNGPAGGVVSPQPHSSYYSGMTGPOHPFYNR 1138

Db 1080 PPLPLHEGPPRAPSGYSQPSVCSSTSFGNGPAGGVVSPQPHSSYYSGMTGPOHPFYNR 1139  
QY 1139 -----RSGSP 1143  
Db 1140 FFAPLYTPRYYPGSGHLSRPSVKTSPLRQDNNGLEVIKEDAAEGLSPSTDSSRSGSP 1199  
QY 1144 AGPVLVLSLNDVACEKIQIEGLDOSMLPOYCTTIKKANINGVLAOCNIDELKKEM 1203  
Db 1200 AGPVLVLSLNDVACEKIQIEGLDOSMLPOYCTTIKKANINGVLAOCNIDELKKEM 1259  
QY 1204 NMNFGDWHLFRSTVLEMRNAESHVPEPRFISESSSGPAPGEPARRASHNELPHTELS 1263  
Db 1260 NMNFGDWHLFRSTVLEMRNAESHVPEPRFISESSSGPAPGEPARRASHNELPHTELS 1319  
QY 1264 SOTPYTLNFSPELNTLGLDEGAPRHSNLSWQSOTRTPSLSLNSQDSSIEISKLTDKV 1323  
Db 1320 SOTPYTLNFSPELNTLGLDEGAPRHSNLSWQSOTRTPSLSLNSQDSSIEISKLTDKV 1379  
QY 1324 QAEYRDAYREYTAQMSOLEGGPGSTTISGRSSPHSTVYMGSSSGSIHNSLPOEKGD 1383  
Db 1380 QAEYRDAYREYTAQMSOLEGGPGSTTISGRSSPHSTVYMGSSSGSIHNSLPOEKGD 1439  
QY 1384 EPKPDGRKSFMLKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLPGKKSSER 1443  
Db 1440 EPKPDGRKSFMLKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLPGKKSSER 1499  
QY 1444 SSLFQTDLKLKSGGLRYQKLPDEDESGTEESDNTPLKDDDKRKAEGKVERVPKSPHS 1503  
Db 1500 SSLFQTDLKLKSGGLRYQKLPDEDESGTEESDNTPLKDDDKRKAEGKVERVPKSPHS 1559  
QY 1504 APPIRTIFAKAYELSDAILDKDSDSGVRSSESSPHSLHNEVADDSOLEKANLIELED 1563  
Db 1560 APPIRTIFAKAYELSDAILDKDSDSGVRSSESSPHSLHNEVADDSOLEKANLIELED 1619  
QY 1564 DSHSGRGIPLHSLSGLODPIIARMSICSEDKSPSCSLIASSPEENWPACOKAYNLNRT 1623  
Db 1620 DSHSGRGIPLHSLSGLODPIIARMSICSEDKSPSCSLIASSPEENWPACOKAYNLNRT 1679  
QY 1624 PSTVTLNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQENLKSMTKRQR 1683  
Db 1680 PSTVTLNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQENLKSMTKRQR 1739  
QY 1684 SSTRLSKDPPHLLHAAASSESTGFGBERESIL 1715  
Db 1740 SSTRLSKDPPHLLHAAASSESTGFGBERESIL 1771

## RESULT 6

US-10-117-229-7  
; Sequence 7, Application US/10117229  
; Publication No. US20030190625A1  
; GENERAL INFORMATION:  
; APPLICANT: OriGene Technologies, Inc.  
; TITLE OF INVENTION: Human Kidins220Pc  
; FILE REFERENCE: 9U 104 R1  
; CURRENT APPLICATION NUMBER: US/10/117,229  
; CURRENT FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 1762  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-117-229-7

Query Match 94.1%; Score 8362.5; DB 14; Length 1762;  
Best Local Similarity 91.6%; Pred. No. 0;  
Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;  
QY 1 MSVLISQSVINYVEENIPALKALLECKDVDERNECCQTPLMAAEOGNLEIVKELIKN 60  
Db 1 MSVLISQSVINYVEENIPALKALLECKDVDERNECCQTPLMAAEOGNLEIVKELIKN 60

QY 61 GANCNLEDLNWTALISAKEGHVIHVBELLKCGVNLEHRDMGGWTALMWACYKGRTDV 120  
DB 61 GANCNLEDLNWTALISAKEGHIHIVBELLKSGASLEHRDMGGWTALMWACYKGRTDV 120  
QY 121 ELLLSHGANSVWGLOYSVYPIIWAAGRHADIHLLLLONGAKVNSDKYGTTPVWAAR 180  
DB 121 ELLLSHGANSVYGL-YSVYPIIWAAGRHADIHLLLLONGAKVNSDKYGTTPVWAAR 179  
QY 181 KGHLECVKHLAMGADVDEGANSMTALIVAKGGYTOSVKEILKRNPNVNLTKDGN 240  
DB 180 KGHLECVKHLAMGADVDEGANSMTALIVAKGGYTOSVKEILKRNPNVNLTKDGN 239  
QY 241 LMIAKESGHEIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 300  
DB 240 LMIAKESGHEIIVQDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQYADIDIRG 299  
QY 301 QDNKTALYWAVEKGNATVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 360  
DB 300 QDNKTALYWAVEKGNATVRDILQCNPDTEICTKDGTEPLIKATMRNIEVVELLDKGA 359  
QY 361 KVSADVKKGDTPLHIAIRGRSKLAELLLRNPDKGRLLYRPNKAGETPYNIDCSHOKSIL 420  
DB 360 KVSADVKKGDTPLHIAIRGRSRLAELLLRNPDKGRLLYRPNKAGETPYNIDCSHOKSIL 419  
QY 421 TQIFGARHLSTETDGMGLYDLYSSALADILSBPTMOPPICVGLYAOQWGSFKFLKCL 480  
DB 420 TQIFGARHLSTETDGMGLYDLYSSALADILSBPTMOPPICVGLYAOQWGSFKFLKCL 479  
QY 481 EDEMKTFAGQIEPLFQPSWLIIVFLTLCCGLGLLFAFTVHNLGIAVLSFLALLYIP 540  
DB 480 EDEMKTFAGQIEPLFQPSWLIIVFLTLCCGLGLLFAFTVHNLGIAVLSFLALLYIP 539  
QY 541 FIVYFGRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPOTTKALPVRLFL 600  
DB 540 FIVYFGRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPOTTKALPVRLFL 599  
QY 601 TDYNRLSVGGETSLAEMIATLSACEREGFLATRLFRVFKTEDTQCKKKWKTCCCLPS 660  
DB 600 TDYNRLSVGGETSLAEMIATLSACEREGFLATRLFRVFKTEESQKKWKWKTCCCLPS 659  
QY 661 FVIFLFIIGCIISGITLLAIIRVDPKHLTVNAVILISASVVGFLAVLNCRTWQVLSLL 720  
DB 660 FVIFLFIIGCIISGITLLAIIRVDPKHLTVNAVILISASVVGFLAVLNCRTWQVLSLL 719  
QY 721 NSQKRLHNAASKLHLKSEGFVKLCEVELMARMAKTIDSTONOTRLVVIIDGDAC 780  
DB 720 NSQKRLHSAASKLHLKSEGFVKLCEVELMARMAKTIDSTONOTRLVVIIDGDAC 779  
QY 781 EQDKVLQMLDTRVRLFSKGPPIAIFASDPHIIKAINQNLNSVLDRSNGHDYMRNIVH 840  
DB 780 EQDKVLQMLDTRVRLFSKGPPIAIFASDPHIIKAINQNLNSVLDRSNGHDYMRNIVH 839  
QY 841 LPVFLNSRGLSNARKFLVTSATNGDPCSDTTGQEDADRRVSONSGEMTKLGSKTALN 900  
DB 840 LPVFLNSRGLSNARKFLVTSATNGDITCSDTTGQEDTDRVSONSLGEMTKLGSKTALN 899  
QY 901 RRDYVRQORTITROMSFDLTKLVTEWFSDISPOTMERLANIVSVTGLLRANOIS 960  
DB 900 RRDYVRQORTITROMSFDLTKLVTEWFSDISPOTMERLANIVSVTGLLRANOIT 959  
QY 961 FNWDLASWNLTEQWPVRTSWLLYLEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020  
DB 960 FNWDLASWNLTEQWPVRTSWLLYLEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1019  
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLRHIIADVRAAREQISIGLAY 1080  
DB 1020 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLRHIIADVRAAREQINIGLAY 1079  
QY 1081 PPLPLHEGPPRPSGYSQPPSCVCSSTFNGPFAGVUSPQPHSSVYSGMTGPQHPFYNR- 1139  
DB 1080 PPLPLHEGPPRPSGYSQPPSCVCSSTFNGPFAGVUSPQPHSSVYSGMTGPQHPFYNR 1139  
QY 1140 -----GSG-----PAGPVL 1151

DB 1140 PFAPVLYTPRYYPGSGQHLISRSSVKTSLSLPRDQNNGLPCDSGFNKQQAAPATGSSLL 1199  
QY 1152 NSLVNDVACEKLOIEGLDQSMLOPYCTTIKANKINGVLAOCNIDELKEMAMNFGDWH 1211  
DB 1200 SSMYTDVVCERQIEGLDQSMMPQYCTTIKANKINGVLSQCNIDELKEMAMNFGDWH 1259  
QY 1212 LFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARRASHNELPHELTELSSQTPYTLN 1271  
DB 1260 LFRSVLEMRSVESQVVPEDPRFLNENSSAPVPHGESARRSHTLPLTELSSQTPYTLN 1319  
QY 1272 PSFELMTGLDEGAPRHSNLSWOSQTRRTTSPSLNSQSSSIETSKLTDKVQAEYRDAY 1331  
DB 1320 PSFELMTGLDEGAPRHSNLSWOSQTRRTTSPSLNSQSSSIETSKLTDKVQAEYRDAY 1379  
QY 1332 REYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGSIHSNLEOEKGDSEPKPDGR 1391  
DB 1380 REYIAQMSQLEGGPGSTTISGRSSPHSTYYIGQSSSGSIHSTLEQERKGEGLQEDGR 1439  
QY 1392 KSFLMKRGDVIDYSSGVSTNDASPLDPIETEDSKSDQSGSKLLPGKSSRSLSLFTD 1451  
DB 1440 KSFLMKRGDVIDYSSGVSTNEASPLDPIETEDSKSDQSGSKLLPGKSSRSLSLFTD 1499  
QY 1452 KLKSGGLRYQKLPSEDESGTEESDNTPLKDDKDKAEGKVERVPKSPHSABPIRTFI 1511  
DB 1500 KLKSGGLRYQKLPSEDESGTEESDNTPLKDDKDKAEGKVERVPKSPHSABPIRTFI 1559  
QY 1512 KAKELYSDALLDKKSDSDSGVRSSSESPHSNLSHNEVADDSDQLEKANLIEDEDDSHSGKRG 1571  
DB 1560 KAKELYSDALLDKKSDSDSGVRSSSESPHSNLSHNEVADDSDQLEKANLIEDEDDSHSGKRG 1619  
QY 1572 IPHLSGLQDPIIARMSICSEDKKSPSECSLIASSPEENWPACOKAYNLNRTPTVTLNN 1631  
DB 1620 MPHLSGLQDPIIARMSICSEDKKSPSECSLIASSPEESWPACOKAYNLNRTPTVTLNN 1679  
QY 1632 NSAPANRANQFDMEGIRETSQVILRPSSSPNPTTIQNLNLSKMTSKRSQSSYTRLSK 1691  
DB 1680 NTAPTNANQFDEBIEGIRETSQVILRPSPNPTAQNLNLSKMAHKRSQSSYTRLSK 1739  
QY 1692 DPELHAAASESTGFGEERESIL 1715  
DB 1740 DASELH-AASESTGFGEERESIL 1762  
RESULT 7  
US-10-021-571-2  
; Sequence 2, Application US/10021571  
; Publication No. US20030166056A1  
; GENERAL INFORMATION:  
; APPLICANT: KONG, Haeyoung  
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN  
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTI  
; FILE REFERENCE: CHAO11A  
; CURRENT APPLICATION NUMBER: US/10/021,571  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/256,909  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1715  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-021-571-2  
Query Match 93.0%; Score 8266; DB 14; Length 1715;  
Best Local Similarity 92.7%; Pred. No. 0;  
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;  
QY 1 MSLVISQSVINYVEENIPALKLEKCKVDNERNECGQTPMLTAASQGNLEIVKELIKN 60  
DB 1 MSLVISQSVINYVEENIPALKLEKCKVDNERNECGQTPMLTAASQGNVIVKELIKN 60



```
QY 61 GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHHRDMGWTALMWACYKGRDVV 120
Db 61 GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHHRDMGWTALMWACYKGRDVV 120
QY 121 ELLSHGANPSVTGLQVSVPIIWAAGRHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
Db 121 ELLSHGANPSVTGLQVSVPIIWAAGRHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
QY 181 KGHLECVKHLAMGADVQEGANSMTALIIVAVRGYVTSQVKEILKRNPNVNLTKDGNNTA 240
Db 181 KGHLECVKHLAMGADVQEGANSMTALIIVAVRGYVTSQVKEILKRNPNVNLTKDGNNTA 240
QY 241 LMTASKEGHEIIVQDLDAGTYNIPDRSGDTVLIIGAVRGHVEIVRALLQKADIDIRG 300
Db 241 LMTASKEGHEIIVQDLDAGTYNIPDRSGDTVLIIGAVRGHVEIVRALLQKADIDIRG 300
QY 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLDKGA 360
Db 301 QDNKTALYWAVEKGNATWVRDILQCNPDTEICTKDGETPLIKATKMRNIEVVELLDKGA 360
QY 361 KVSADVKKGDTPLHVAIRGRSRKLAELLRNPKDGRLLYRPNKAGETPPYINDCSHQKSIL 420
Db 361 KVSADVKKGDTPLHVAIRGRSRKLAELLRNPKDGRLLYRPNKAGETPPYINDCSHQKSIL 420
QY 421 TQIFGARHLSPTETDGMGLDYLYSSALADILSEPTWOPPICVGLYAQWGSFKLLKKL 480
Db 421 TQIFGARHLSPTETDGMGLDYLYSSALADILSEPTWOPPICVGLYAQWGSFKLLKKL 480
QY 481 EDEKTFAGQIQBPLOFQSWLIIVFLTLCCGGLGFAFVPHNLGTAVALSLFALLIYF 540
Db 481 EDEKTFAGQIQBPLOFQSWLIIVFLTLCCGGLGFAFVPHNLGTAVALSLFALLIYF 540
QY 541 FIVIFGRRGEGESWNMAWVLSRLARHIGYLELLKLMFVNPPPELPEQTTKALPVRF 600
Db 541 FIVIFGRRGEGESWNMAWVLSRLARHIGYLELLKLMFVNPPPELPEQTTKALPVRF 600
QY 601 TDYNRLSSVGGETSLAMIAITLSDACERERGFATRLFRVFKTEDTGKKKWKTCCLPS 660
Db 601 TDYNRLSSVGGETSLAMIAITLSDACERERGFATRLFRVFKTEDTGKKKWKTCCLPS 660
QY 661 FVIFLFTIGCIISGITLTAIFRVDPKHLTVNAVLIISIASVVGAFVLCNRTWQVLDL 720
Db 661 FVIFLFTIGCIISGITLTAIFRVDPKHLTVNAVLIISIASVVGAFVLCNRTWQVLDL 720
QY 721 NSQKRLLHNAASKLHLKLSGFMKVLKCEVELMARMAKTTIDSFTQNTQLRWIIDGLDAC 780
Db 721 NSQKRLLHNAASKLHLKLSGFMKVLKCEVELMARMAKTTIDSFTQNTQLRWIIDGLDAC 780
QY 781 EQDKVLQMLDTRVLFSGKGFPIAFASDPHIIKAINONLNSVLNDSNINCHDYMNRIVH 840
Db 781 EQDKVLQMLDTRVLFSGKGFPIAFASDPHIIKAINONLNSVLNDSNINCHDYMNRIVH 840
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPCTDTTGQEDADRRVSONSLGEMTKLGSKTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPCTDTTGQEDADRRVSONSLGEMTKLGSKTALN 900
QY 901 RRDYTRRQMTITRQMSFDLTKLVTEDWFSDISPQTMRRLLNIIVSVTGRLLRANQIS 960
Db 901 RRDYTRRQMTITRQMSFDLTKLVTEDWFSDISPQTMRRLLNIIVSVTGRLLRANQIT 960
QY 961 FNMDRLASWINLTQWPRYSWLLIYLBETEGIPDQMTLKTIVYRISKNIPTTKDVEPLL 1020
Db 961 FNMDRLASWINLTQWPRYSWLLIYLBETEGIPDQMTLKTIVYRISKNIPTTKDVEPLL 1020
QY 1021 EIDGDIENFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGLAY 1080
Db 1021 EIDGDIENFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQINIGLAY 1080
QY 1081 PPLPLHEGPPRAPSGYQPPSVCSSTSFNGFPAGVSVSPQPHSGYSYSGMTGPQHPFYNRG 1140
Db 1081 PPLPLHEGPPRAPSGYQPPSVCSSTSFNGFPAGVSVSPQPHSGYSYSGMTGPQHPFYNRG 1140
```

```
QY 1141 SGAPGPPVLLNSLNDVAVCEKLOIEGLDQSMLOPOYCTTIKKANINGRVLACNIDELK 1200
Db 1141 AVPATGSSLLLSMTVDVCEKLOIEGLDQSMLOPOYCTTIKKANINGRVLACNIDELK 1200
QY 1201 KEMNNMFGDWHLFRSTVLEMRNABSHVVPEDPRFLSESSSGPAPHGEARRASHNELPHT 1260
Db 1201 KEMNNMFGDWHLFRSVLEMRNABSHVVPEDPRFLSESSSGPAPHGEARRASHNELPHT 1260
QY 1261 ELSQTPYTLNPSPELNTLGLDEGAPRHSNLSWQSTRTPTSLSLNSQDSSIEISKL 1320
Db 1261 ELSQTPYTLNPSPELNTLGLDEGAPRHSNLSWQSTRTPTSLSLNSQDSSIEISKL 1320
QY 1321 DKVQAYRDAYREYIAQMSQLEGCGSTTISGRSSPHSTVYMGQSSGSGSIHNLQOEKG 1380
Db 1321 DKVQAYRDAYREYIAQMSQLEGCGSTTISGRSSPHSTVYMGQSSGSGSIHNLQOEKG 1380
QY 1381 KQSEKPPDDGRKSFMLKRGDVIDYSSGVSTNDASPLDPIETEDEKSDQSGSKLLPKKS 1440
Db 1381 KQSEKPPDDGRKSFMLKRGDVIDYSSGVSTNDASPLDPIETEDEKSDQSGSKLLPKKS 1440
QY 1441 SERSSLFQDLDKLGSLRYQKLPSEDESGTSEESNTPLKDDKDKRKAEGKVERVPKS 1499
Db 1441 SERSSLFQDLDKLGSLRYQKLPSEDESGTSEESNTPLKDDKDKRKAEGKVERVPKS 1499
QY 1500 PEHSAPIRTFIKAKYLSDALDKKSDSDSGVRSSESPNHSNHLNHEVADDSOLEKANLI 1559
Db 1500 PEHSAPIRTFIKAKYLSDALDKKSDSDSGVRSSESPNHSNHLNHEVADDSOLEKANLI 1559
QY 1561 ELEDHSHGKRGPHSLGSLQDPIIARMSICSEDKSPSECSLIASSPEENWPACQKAYN 1619
Db 1561 ELEDHSHGKRGPHSLGSLQDPIIARMSICSEDKSPSECSLIASSPEENWPACQKAYN 1619
QY 1620 LNRTPTSTVLNNSAPANRANQNFDEMEGIRETSQVILRPPSSPNPTTQENLKSMTXK 1679
Db 1620 LNRTPTSTVLNNSAPANRANQNFDEMEGIRETSQVILRPPSSPNPTTQENLKSMTXK 1679
QY 1680 RSQRSSYTRLSKDPPELHAAASSESTGFGEERESIL 1715
Db 1680 RSQRSSYTRLSKDPPELHAAASSESTGFGEERESIL 1715
QY 1681 RSQRSSYTRLSKDPPELHAAASSESTGFGEERESIL 1715
Db 1681 RSQRSSYTRLSKDPPELHAAASSESTGFGEERESIL 1715
```

## RESULT 8

```
US-10-117-229-6
; Sequence 6, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220Pc
; FILE REFERENCE: 9U 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 6
; LENGTH: 1715
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-117-229-6
```

```
Query Match 93.0%; Score 8266; DB 14; Length 1715;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 1591; Conservative 53; Mismatches 70; Indels 2; Gaps 2;
```

```
QY 1 MSLVLSQSIVNYVEENIPALKALLEKCKVDNRNECGQTPLMIAAFOGNEIVKELIKN 60
Db 1 MSLVLSQSIVNYVEENIPALKALLEKCKVDNRNECGQTPLMIAAFOGNEIVKELIKN 60
QY 61 GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHHRDMGWTALMWACYKGRDVV 120
Db 61 GANCNLEDLNDWTALISASKEGHVHIVEELLKCGVNLHHRDMGWTALMWACYKGRDVV 120
QY 121 ELLSHGANPSVTGLQVSVPIIWAAGRHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
Db 121 ELLSHGANPSVTGLQVSVPIIWAAGRHADIVHLLLQNGAKVNSDKYGTTPLVWAAR 180
```



Db 121 ELLSHGANPSVTGLQYVYPIIIWAAGRGHADIHVHLLONGAKVNCSDKYGTTPVJVAAR 180  
Qy 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA 240  
Db 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA 240  
Qy 241 LMIASKEGHEIVQDLDAGTYVNIIPDRSGDTVLIVAVRGHVEIVRALLQYADIDIRG 300  
Db 241 LMIASKEGHEIVQDLDAGTYVNIIPDRSGDTVLIVAVRGHVEIVRALLQYADIDIRG 300  
Qy 301 QDNKTALYVAVEKGNATVRDILQCNPDTEICTKDGETPLIKATKORNIUEVLELLDKGA 360  
Db 301 QDNKTALYVAVEKGNATVRDILQCNPDTEICTKDGETPLIKATKORNIUEVLELLDKGA 360  
Qy 361 KVSADVKKGDTPHIAIRGRSKLAELLRLNPKGRILLYRNPKAGETPNYIDCSHOKSIL 420  
Db 361 KVSADVKKGDTPHIAIRGRSKLAELLRLNPKGRILLYRNPKAGETPNYIDCSHOKSIL 420  
Qy 421 TOIFGARHLSPTETDGMGLYDLYSSALADILSEPTMOPPICVGLYAWGSGKFLFKKL 480  
Db 421 TOIFGARHLSPTETDGMGLYDLYSSALADILSEPTMOPPICVGLYAWGSGKFLFKKL 480  
Qy 481 EDEMKTFAGQOEPLFQPSWLIIVELTLLCGGLLPAFTVHPNIGIAVSLFALLIYIF 540  
Db 481 EDEMKTFAGQOEPLFQPSWLIIVELTLLCGGLLPAFTVHPNIGIAVSLFALLIYIF 540  
Qy 541 FIVYFGRRGESWNWAWLSTRLARHIGYLELLKLMFVNPPELPOTTKALPVRFPLF 600  
Db 541 FIVYFGRRGESWNWAWLSTRLARHIGYLELLKLMFVNPPELPOTTKALPVRFPLF 600  
Qy 601 TDYNRLSSVGGTSLAEMIATLSACEREFGLATRLFRVPKTEDTQKKWKKTCCILPS 660  
Db 601 TDYNRLSSVGGTSLAEMIATLSACEREFGLATRLFRVPKTEDTQKKWKKTCCILPS 660  
Qy 720 FVIFLFIIGCIIIGITLLAIIRVDPKHTLVNAVLIISASVVGGLAPVNLNCRTWQVLSLL 720  
Db 720 FVIFLFIIGCIIIGITLLAIIRVDPKHTLVNAVLIISASVVGGLAPVNLNCRTWQVLSLL 720  
Qy 721 NSQKRLHNAASKLHLKSEGFVKLVCEVELMARMATIDSFONQTRLVVITDGLDAC 780  
Db 721 NSQKRLHNAASKLHLKSEGFVKLVCEVELMARMATIDSFONQTRLVVITDGLDAC 780  
Qy 781 EODKVLQMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLNDSNINGHDYMRNVH 840  
Db 781 EODKVLQMLDTRVRLFSKGPPIAFASDPHIIKAINQNLSVLNDSNINGHDYMRNVH 840  
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDVCSDTGTQEDADRVSONSLGEMTKLASKTALN 900  
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVCSDTGTQEDADRVSONSLGEMTKLASKTALN 900  
Qy 901 RRDYVRRQOMORTITROMSPDLTKLVTEDWFSDISPOTWRLANIVSVTGRLARANQIS 960  
Db 901 RRDYVRRQOMORTITROMSPDLTKLVTEDWFSDISPOTWRLANIVSVTGRLARANQIS 960  
Qy 961 FNWDLASWINLTQWPPVTSWLIYLEETEIGIPDQMTLKIYERISKNIPPTTKDVEPLL 1020  
Db 961 FNWDLASWINLTQWPPVTSWLIYLEETEIGIPDQMTLKIYERISKNIPPTTKDVEPLL 1020  
Qy 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLRIBIIADVRAAREQINIGLAY 1080  
Db 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLRIBIIADVRAAREQINIGLAY 1080  
Qy 1081 PPLPLHGGPPRPPSGYQSPSCVSTSPNGPAGVVSQPHSSYYSGMTQPOHFFNVRG 1140  
Db 1081 PPLPLHGGPPRPPSGYQSPSCVSTSPNGPAGVVSQPHSSYYSGMTQPOHFFNVRG 1140  
Qy 1141 SGAPGPPVLIANSVNDVAVCEKLOIEGLDOSMLPOVCTTIKKANINGRVLQACNIDELK 1200  
Db 1141 AVPATGSLLSMTVDVCEKLRQIEGLDQNMVPOVCTTIKKANINGRVLQACNIDELK 1200  
Qy 1201 KEMANNFQDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGPARRASHNELPHT 1260  
Db 1201 KEMANNFQDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPHGPARRASHNELPHT 1260

Qy 1261 ELSSQTPYTLNFSPEELNTLGLDEGAPRHSLNLSWQSQTRRTTPSLSSLSNQSSSIISKLT 1320  
Db 1261 ELSSQTPYTLNFSPEELNTLGLDEGAPRHSLNLSWQSQTRRTTPSLSSLSNQSSSIISKLT 1320  
Qy 1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSPSHSTYYMGSSSGSIHNSLEQKRG 1380  
Db 1321 DKVQAEYRDAYREYIAQMSQLEGGPGSTTISGRSPSHSTYYMGSSSGSIHNSLEQKRG 1380  
Qy 1381 KDSPEKPDGKSKFLMKRGDVIDYSSSGVSTNDASPLDPIITEEDEKSDQSGSKLLPGKKS 1440  
Db 1381 KEGELKQEDGKSKFLMKRGDVIDYSSSGVSTNDASPLDPIITEEDEKSDQSGSKLLPGKKS 1440  
Qy 1441 SERSLFOTDLKLGSGRLYOKLSEDESTEESDNTPLKDDKOKRKAQKGVKERVPKS 1499  
Db 1441 SERSLFOTDLKLGSGRLYOKLSEDESTEESDNTPLKDDKOKRKAQKGVKERVPKS 1499  
Qy 1500 PEHSAEPIRTPIKAKYLSGLQDPIIARMSICSDKSPSECSLIASSPEENWPACQKAYN 1559  
Db 1501 QEHSAEPIRTPIKAKYLSGLQDPIIARMSICSDKSPSECSLIASSPEENWPACQKAYN 1559  
Qy 1560 ELEDSSHGKRGIPHLSGLQDPIIARMSICSDKSPSECSLIASSPEENWPACQKAYN 1619  
Db 1561 ELEDSSHGKRGIPHLSGLQDPIIARMSICSDKSPSECSLIASSPEENWPACQKAYN 1619  
Qy 1620 LNRTPTVTLLNNSAPANRANONFDEMEGIRETSQVILRPPSSPNPTTIONENLKSMTWK 1679  
Db 1621 LNRTPTVTLLNNSAPANRANONFDEMEGIRETSQVILRPPSSPNPTTIONENLKSMTWK 1679  
Qy 1680 RSQRSSYTRLKSKOPPELHAAASSESTGFEERESIL 1715  
Db 1681 RSQRSSYTRLKSKOPPELHAAASSESTGFEERESIL 1715

## RESULT 9

US-10-117-229-3  
; Sequence 3, Application US/10117229  
; Publication No. US20030190625A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc.  
; TITLE OF INVENTION: Human Kidins220PC  
; FILE REFERENCE: 9U 104 R1  
; CURRENT APPLICATION NUMBER: US/10/117,229  
; CURRENT FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 3  
; LENGTH: 1184  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-117-229-3

Query Match 66.4%; Score 5899.5; DB 14; Length 1184;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1138; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MSVLISQSVINYYVEENIPALKALLEKCKVDNERNECGQTPLMIAAECQNLIEIVKELIKN 60  
Db 1 MSVLISQSVINYYVEENIPALKALLEKCKVDNERNECGQTPLMIAAECQNLIEIVKELIKN 60  
Qy 61 GANCNLEDLONWTALISASKEGHVHIVEELLKCVNLHHRDMGGWTALMWACYKGRDVT 120  
Db 61 GANCNLEDLONWTALISASKEGHVHIVEELLKCVNLHHRDMGGWTALMWACYKGRDVT 120  
Qy 121 ELLLSHGANSVTLQYQSVYPIIIWAAGRGHADIHVHLLONGAKVNCSDKYGTTPVJVAAR 180  
Db 121 ELLLSHGANSVTLQYQSVYPIIIWAAGRGHADIHVHLLONGAKVNCSDKYGTTPVJVAAR 180  
Qy 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA 240  
Db 181 KGHLECVKHLAMGADVQEGANSMTALIVAVKGGYTQSVKEILKRNPNVNLTDKGNNTA 240  
Qy 241 LMIASKEGHEIVQDLDAGTYVNIIPDRSGDTVLIVAVRGHVEIVRALLQYADIDIRG 300

```
Db 240 LMIASKEGHEITEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
QY 301 QDNKTALYWAVEKGNATWVRDILQCPNDEICTKDGSETPLIKATKMRNIEVVELLDKGA 360
Db 300 QDNKTALYWAVEKGNATWVRDILQCPNDEICTKDGSETPLIKATKMRNIEVVELLDKGA 359
QY 361 KVSADVKKGDTPLHIAIRGSRKLAELLNRPKDGRLLYRPNKAGETPPYINDCSHOKSIL 420
Db 360 KVSADVKKGDTPLHIAIRGSRKLAELLNRPKDGRLLYRPNKAGETPPYINDCSHOKSIL 419
QY 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPIICVGLYAQWGSFGSKFLKKL 480
Db 420 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPIICVGLYAQWGSFGSKFLKKL 479
QY 481 EDEMKTFAQQIIEPLFQFSLWLVFLTLCCGGLGLLFAFTVHPNLGIAVSLFALLIYIF 540
Db 480 EDEMKTFAQQIIEPLFQFSLWLVFLTLCCGGLGLLFAFTVHPNLGIAVSLFALLIYIF 539
QY 541 FIVIFGGRREGESWNAWVLSLRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 600
Db 540 FIVIFGGRREGESWNAWVLSLRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 599
QY 601 TDYNRSLSSVGGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKKTCCPLS 660
Db 600 TDYNRSLSSVGGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKKTCCPLS 659
QY 661 FVIFLFIIGCIISGITLLAIIRVDPKHLTVNAVLISIASVVGGLAFVNLNCRWTWQVLDL 720
Db 660 FVIFLFIIGCIISGITLLAIIRVDPKHLTVNAVLISIASVVGGLAFVNLNCRWTWQVLDL 719
QY 721 NSQRKELHNAASKHLKSGEFGMKVLKCEVELMARMAKTIDSFTQNTQRLVWIIDGLDAC 780
Db 720 NSQRKELHNAASKHLKSGEFGMKVLKCEVELMARMAKTIDSFTQNTQRLVWIIDGLDAC 779
QY 781 EQDKVLOMLDTRVRLFSKGFPIAIFASDPHIIKAINQNLNSVLRDSNINGHDMYRNIVH 840
Db 780 EQDKVLOMLDTRVRLFSKGFPIAIFASDPHIIKAINQNLNSVLRDSNINGHDMYRNIVH 839
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDITQIEDADRRVSQNSLGMTKLGSKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDVPSCDITQIEDADRRVSQNSLGMTKLGSKTALN 899
QY 901 RRDYRRRQWQRTITRQMSFDLTCLIVTEWDFSDISFQTVRRLLNIVSVTGRLLRANQIS 960
Db 900 RRDYRRRQWQRTITRQMSFDLTCLIVTEWDFSDISFQTVRRLLNIVSVTGRLLRANQIS 959
QY 961 FNNDRLASWNLTEQWPYRTSWILYLEETEGIPDQWTLKTIYERISKNIPTTKDVEPLL 1020
Db 960 FNNDRLASWNLTEQWPYRTSWILYLEETEGIPDQWTLKTIYERISKNIPTTKDVEPLL 1019
QY 1021 EIDGDIRNFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIADVRAAREQISIGLAY 1080
Db 1020 EIDGDIRNFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIADVRAAREQISIGLAY 1079
QY 1081 PPLPLHEGPPRAPSGVSQPSVCSSTSFNGPPAGGVSPPOPHSSYSGMTGPOHPFYNR 1139
Db 1080 PPLPLHEGPPRAPSGVSQPSVCSSTSFNGPPAGGVSPPOPHSSYSGMTGPOHPFYNR 1138
```

## RESULT 10

```
US-10-117-229-9
; Sequence 9, Application US/10117229
; Publication No. US20030190625A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.
; TITLE OF INVENTION: Human Kidins220PC
; FILE REFERENCE: 90 104 R1
; CURRENT APPLICATION NUMBER: US/10/117,229
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
```

```
; LENGTH: 1184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-229-9
```

## Query Match

Best Local Similarity 66.4%; Score 5899.5; DB 14; Length 1184;

Mismatches 1138; Conservative 0; Indels 1; Gaps 1;

```
QY 1 MSVLISSQSVINVEENIIPALKALLEKCKDVBERNECGQTPLMIAAEOGNLEIVKELIKN 60
Db 1 MSVLISSQSVINVEENIIPALKALLEKCKDVBERNECGQTPLMIAAEOGNLEIVKELIKN 60
QY 61 GANCNLEDNNTALISASKEGHVHIVEELLKCGVNLHRDMGGWTALWACVKGRTDVV 120
Db 61 GANCNLEDNNTALISASKEGHVHIVEELLKCGVNLHRDMGGWTALWACVKGRTDVV 120
QY 121 ELLSHGANPSVTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLVWAAR 180
Db 121 ELLSHGANPSVTGLQYSVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLVWAAR 179
QY 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILLKRNPNVNLTDKQNTA 240
Db 180 KGHLECVKHLAMGADVDOEGANSMTALIVAKGGYTQSVKEILLKRNPNVNLTDKQNTA 239
QY 241 LMIASKEGHEITEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 300
Db 240 LMIASKEGHEITEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRG 299
QY 301 QDNKTALYWAVEKGNATWVRDILQCPNDEICTKDGSETPLIKATKMRNIEVVELLDKGA 360
Db 300 QDNKTALYWAVEKGNATWVRDILQCPNDEICTKDGSETPLIKATKMRNIEVVELLDKGA 359
QY 361 KVSADVKKGDTPLHIAIRGSRKLAELLNRPKDGRLLYRPNKAGETPPYINDCSHOKSIL 420
Db 360 KVSADVKKGDTPLHIAIRGSRKLAELLNRPKDGRLLYRPNKAGETPPYINDCSHOKSIL 419
QY 421 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPIICVGLYAQWGSFGSKFLKKL 480
Db 420 TQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPIICVGLYAQWGSFGSKFLKKL 479
QY 481 EDEMKTFAQQIIEPLFQFSLWLVFLTLCCGGLGLLFAFTVHPNLGIAVSLFALLIYIF 540
Db 480 EDEMKTFAQQIIEPLFQFSLWLVFLTLCCGGLGLLFAFTVHPNLGIAVSLFALLIYIF 539
QY 541 FIVIFGGRREGESWNAWVLSLRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 600
Db 540 FIVIFGGRREGESWNAWVLSLRLARHIGYLELLKLMFVNPPELPEQTTKALPVRFLF 599
QY 601 TDYNRSLSSVGGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKKTCCPLS 660
Db 600 TDYNRSLSSVGGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQGGKKWKKTCCPLS 659
QY 661 FVIFLFIIGCIISGITLLAIIRVDPKHLTVNAVLISIASVVGGLAFVNLNCRWTWQVLDL 720
Db 660 FVIFLFIIGCIISGITLLAIIRVDPKHLTVNAVLISIASVVGGLAFVNLNCRWTWQVLDL 719
QY 721 NSQRKELHNAASKHLKSGEFGMKVLKCEVELMARMAKTIDSFTQNTQRLVWIIDGLDAC 780
Db 720 NSQRKELHNAASKHLKSGEFGMKVLKCEVELMARMAKTIDSFTQNTQRLVWIIDGLDAC 779
QY 781 EQDKVLOMLDTRVRLFSKGFPIAIFASDPHIIKAINQNLNSVLRDSNINGHDMYRNIVH 840
Db 780 EQDKVLOMLDTRVRLFSKGFPIAIFASDPHIIKAINQNLNSVLRDSNINGHDMYRNIVH 839
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDITQIEDADRRVSQNSLGMTKLGSKTALN 900
Db 840 LPVFLNSRGLSNARKFLVTSATNGDVPSCDITQIEDADRRVSQNSLGMTKLGSKTALN 899
QY 901 RRDYRRRQWQRTITRQMSFDLTCLIVTEWDFSDISFQTVRRLLNIVSVTGRLLRANQIS 960
Db 900 RRDYRRRQWQRTITRQMSFDLTCLIVTEWDFSDISFQTVRRLLNIVSVTGRLLRANQIS 959
```

Qy 961 FNDRLASWNLTEQWVRTSWLILYLBETEGIPDQMTLKTIERISKNIPTTKDVEPLL 1020  
Db 960 FNDRLASWNLTEQWVRTSWLILYLBETEGIPDQMTLKTIERISKNIPTTKDVEPLL 1019  
Qy 1021 EIDGDIRNEFVLSRTPLVARDVKVPLPCTVNLDPKRLRIIADVRAAREQISIGGLAY 1080  
Db 1020 EIDGDIRNEFVLSRTPLVARDVKVPLPCTVNLDPKRLRIIADVRAAREQISIGGLAY 1079  
Qy 1081 PPLPLHEGPPRAPSQYSPSPSCSTSPNGPAGVSPQPHSSYSYSGMTGPQHPFYNR 1139  
Db 1080 PPLPLHEGPPRAPSQYSPSPSCSTSPNGPAGVSPQPHSSYSYSGMTGPQHPFYNR 1138

RESULT 11  
US-10-106-698-6378  
; Sequence 6378, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6378  
; LENGTH: 705  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (244)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: MISC FEATURE  
; LOCATION: (337)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-6378

Query Match 34.3%; Score 3050.5; DB 14; Length 705;  
Best Local Similarity 98.8%; Pred. No. 8.9e-192;  
Matches 587; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 62 ANCNLELDNNWALISASKEGHVHVEELLKCGVNLHRDGGWGTALMWACYKGRTDVVE 121  
Db 28 SNCNLELDNNWALISASKEGHVHVEELLKCGVNLHRDGGWGTALMWACYKGRTDVVE 87  
Qy 122 LLLSHGANPSVTGLQYSVYPIIWAAGRCHADIVHLLQNGAKVNCSDKYGTTPLVWAARK 181  
Db 88 LLLSHGANPSVTGL-QYSVYPIIWAAGRCHADIVHLLQNGAKVNCSDKYGTTPLVWAARK 146  
Qy 182 GHLECVKHLANGADVQEGANSMTALIVAKGGYTSVKELKRNPNVNLTDKDGNTAL 241  
Db 147 GHLECVKHLANGADVQEGANSMTALIVAKGGYTSVKELKRNPNVNLTDKDGNTAL 206  
Qy 242 MIASKEGHEITVDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQ 301  
Db 207 MIASKEGHEITVDLLDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQ 266  
Qy 302 DNKTALYWAKEGNATWVRDILQCNPDTEICTDGETPLIKATKORNIIEVVELLDKGA 361  
Db 267 DNKTALYWAKEGNATWVRDILQCNPDTEICTDGETPLIKATKORNIIEVVELLDKGA 326  
Qy 362 VSAVDKKGDTPLHAIAGRSRKLAEILLRNPKDGRLLYRPNKAGETPNIDCSHQSILIT 421  
Db 327 VSAVDKKGDTPLHAIAGRSRKLAEILLRNPKDGRLLYRPNKAGETPNIDCSHQSILIT 386  
Qy 422 QIFGARHLSPTTGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKGSFLKKLE 481

Db 387 QIFGARHLSPTTGDMLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKGSFLKKLE 446  
Qy 482 DEMKTEAQOIEPLFQPSWLIIVFLTLICCGGLIFLFAFTHVHPNLGIAVLSFLALLYIFF 541  
Db 447 DEMKTEAQOIEPLFQPSWLIIVFLTLICCGGLIFLFAFTHVHPNLGIAVLSFLALLYIFF 506  
Qy 542 IVIYFGRRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPPEQTITKALPVRFLFT 601  
Db 507 IVIYFGRRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPPEQTITKALPVRFLFT 566  
Qy 602 DYNRLSSVGGTSLAEMIATLSDCEREFGLATRLPRVFKTDTQGGKKKWK 655  
Db 567 DYNRLSSVGGTSLAEMIATLSDCEREFGLATRLPRVFKTDTQGGKKKWK 620

RESULT 12  
US-09-835-788A-17  
; Sequence 17, Application US/09835788A  
; Patent No. US20020077458A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a  
; FILE REFERENCE: PT018P1  
; CURRENT APPLICATION NUMBER: US/09/835,788A  
; PRIOR FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: PCT/US00/28666  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/159,585  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 60/167,246  
; PRIOR FILING DATE: 1999-11-24  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 551  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-835-788A-17

Query Match 32.1%; Score 2851.5; DB 9; Length 551;  
Best Local Similarity 99.6%; Pred. No. 7.5e-179;  
Matches 550; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 102 MGGTALMWACYKGRTDVVELLSHGANPSVTGLQYSVYPIIWAAGRCHADIVHLLQNG 161  
Db 1 MGGTALMWACYKGRTDVVELLSHGANPSVTGL-QYSVYPIIWAAGRCHADIVHLLQNG 59  
Qy 162 AKVNCSDKYGTTPLVWAARKGHLECVKHLANGADVQEGANSMTALIVAKGGYTSVK 221  
Db 60 AKVNCSDKYGTTPLVWAARKGHLECVKHLANGADVQEGANSMTALIVAKGGYTSVK 119  
Qy 222 EILKRNPNVNLTDKGNATLMIASKEGHEITVDLLDAGTYVNIIPDRSGDTVLIGAVRG 281  
Db 120 EILKRNPNVNLTDKGNATLMIASKEGHEITVDLLDAGTYVNIIPDRSGDTVLIGAVRG 179  
Qy 282 HVEIVRALLQKYADIDIRGQDNKTALYWAKEGNATWVRDILQCNPDTEICTDGETPLI 341  
Db 180 HVEIVRALLQKYADIDIRGQDNKTALYWAKEGNATWVRDILQCNPDTEICTDGETPLI 239  
Qy 342 KATKORNIIEVVELLDKGAQVSAVDKKGDTPLHAIAGRSRKLAEILLRNPKDGRLLYRP 401  
Db 240 KATKORNIIEVVELLDKGAQVSAVDKKGDTPLHAIAGRSRKLAEILLRNPKDGRLLYRP 299  
Qy 402 NKAGETPNIDCSHQSILITQIFGARHLSPTTGDMLGYDLYSSALADILSEPTMQPPI 461  
Db 300 NKAGETPNIDCSHQSILITQIFGARHLSPTTGDMLGYDLYSSALADILSEPTMQPPI 359  
Qy 462 CVGLYAQWGSKGSFLKKLEDEMKTTFAGQOIEPLFQPSWLIIVFLTLICCGGLIFLFAFTV 521  
Db 360 CVGLYAQWGSKGSFLKKLEDEMKTTFAGQOIEPLFQPSWLIIVFLTLICCGGLIFLFAFTV 419



```

Qy 1563 DDSHSGRGIGPHSLSGLQDPIIARMSICSEDKKSPPSCSLIASPEENWPACQAYNLNR 1622
Db 361 DDSHSGRGIGPHSLSGLQDPIIARMSICSEDKKSPPSCSLIASPEENWPACQAYNLNR 420
Qy 1623 TPSTVTILNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNELKSMWTKRSQ 1682
Db 421 TPSTVTILNNNSAPANRANQNFDEMEGIRETSQVILRPSSSPNPTTIQNELKSMWTKRSQ 480
Qy 1683 RSSYTRLSKOPPELHAAAASSESTGFGEERESIL 1715
Db 481 RSSYTRLSKOPPELHAAAASSESTGFGEERESIL 513

RESULT 15
US-10-021-571-8
; Sequence 8, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, HaeYoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTIB
; FILE REFERENCE: CHA011A
; CURRENT APPLICATION NUMBER: US/10/021,571
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1498
; TYPE: PRT
; ORGANISM: Drosophila
US-10-021-571-8

Query Match 26.8%; Score 2379,5; DB 14; Length 1498;
Best Local Similarity 37.6%; Pred. No. 4,5e-147;
Matches 559; Conservative 259; Mismatches 490; Indels 179; Gaps 39;

Qy 1 MSVLISQSVINYVEEENIPALKALLE-KCKDVPDERNECGQTPLMIAAEQGNLEIVKELIK 59
Db 23 MGSILGHRALLQYIDNNDISGLRAILDSRLHTIDDRDENATTVLMVWVAGRLGTAFAVREFLA 82
Qy 60 NGANCNLEDLDNWTALISASKEGHVHIVEALLKCGVNLHHRDMGCVTALMWACYKGRITDV 119
Db 83 RGADVQRAEDLDNWTALLCASRNGHLDVQVLLDHGAEVEHRDMGCVTSLMWAAYRGHTTEL 142
Qy 120 VELLSHGANPSTVGLQYSVYPIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAA 179
Db 143 VRLLLDKGADGNAHG-NYHLGALLWAAGRYKDIIVELLVQRGAQVNVGDKYGTALVWAC 201
Qy 180 RKGHLECVKHLLAWGADVDDQEGANSMTALIVAKGYGTSQVKEILKRNPNVNLTDXGNT 239
Db 202 RRGNVETVDTLLKAGANVDTAGMYSWTPLLVAAAAGGHTDCVSSILEKKNVNALDXGNT 261
Qy 240 ALMTASKEGHTETVQDLLDAGTVYNIIPDRSGDTVLICAVRGGHVEIVRALLOKYADIDIR 299
Db 262 ALCIASREGQDIAASLIAGAAYINIQDRGADTPLIHAVKAGHRTVVEALLKKHADVDIQ 321
Qy 300 GQDNKTALYWAVEKGNATMVRDILQCNPDPEI CTXKDGETPLIKATKRNIEVVVELLDKX 359
Db 322 GKDKRTAIYTAVEKGHTPIVKLLLATNPDESATKGDTPLLRAVRNRLIEIVHLLDRK 381
Qy 360 AKYSADVKKGDTPLHIAIRGSRKLAELLRNPKOGRLLYRPNKAGETPNYDCSHQKSI 419
Db 382 AKVTASDKRGDTCLIHAMRARSRTIVEALLRNPKHSQLLYRANKAGETPNYIDSLHQKTI 441
Qy 420 LTOIFGARHLSPPTETGCDMLGYDLYSSALADILSEPTMOPPICVGLYAOWGSKSFLKK 479
Db 442 LGQVFGARRLNTNEDSEGMLGYELYSALADVLUSEPTLTPTITVGLYAKWGSKSFLLNK 501
Qy 480 LEDEMTKFAQQIPELPQFQSWLIVFTL-----LLCG-----GLGLLFAFTVHPNIGIAVSLGF 533

```

Search completed: February 5, 2005, 19:54:55  
Job time : 147 secs

Job time : 147 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2005, 19:16:39 ; Search time 31 Seconds  
(without alignments)  
5322.960 Million cell updates/sec

Title: US-10-021-571-4

Perfect score: 8884

Sequence: 1 MSVLISQSVINVEENIPAL.....LHAASSESTGFEERESTL 1715

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5205.5	58.6	1031	T43458	hypothetical prote
2	1832.5	20.6	1398	T21884	hypothetical prote
3	1255	14.1	246	T46392	hypothetical prote
4	555.5	6.3	1881	1 SJHUK	ankyrin 1, erythro
5	552.5	6.2	1856	2 B35049	ankyrin 1, erythro
6	552.5	6.2	1880	2 A35049	ankyrin 1, erythro
7	544	6.1	426	2 AE2149	hypothetical prote
8	539	6.1	3924	2 S37431	ankyrin 2, neurona
9	530	6.0	1961	2 T42716	ankyrin 3, splice
10	529.5	6.0	1940	2 T42715	ankyrin 3, splice
11	525	5.9	1862	2 I49502	ankyrin - mouse
12	523	5.9	1943	2 T42713	ankyrin 3, splice
13	521	5.9	4377	2 A55575	ankyrin 3, long sp
14	516	5.8	1765	2 S37714	ankyrin 3, splice
15	506	5.7	1848	2 S37711	ankyrin, erythrocy
16	475.5	5.4	1549	2 T13940	ankyrin - fruit fl
17	442	5.0	2039	2 T15347	ankyrin-related un
18	415	4.7	2584	2 T24158	hypothetical prote
19	415	4.7	2606	2 T24157	hypothetical prote
20	409.5	4.6	791	2 T42691	hypothetical prote
21	409.5	4.6	1423	1 I37275	death-associated p
22	387.5	4.4	1062	2 T30255	inversin - mouse
23	375.5	4.2	323	2 B47169	ankyrin-like repea
24	373.5	4.2	557	2 T46507	hypothetical prote
25	373.5	4.2	1062	2 T14151	Inv protein - mous
26	349.5	3.9	397	2 T46445	hypothetical prote
27	348	3.9	679	2 A45771	2-5A-dependent RNA
28	343	3.9	741	2 A45771	2-5A-dependent RNA
29	336.5	3.8	991	2 T25412	hypothetical prote

30	329	3.7	1188	2 T19552	hypothetical prote
31	328	3.7	1401	2 S11527	alpha-latrotoxin p
32	325	3.7	934	1 H71274	probable ankyrin
33	320.5	3.6	441	2 S74626	erythroid ankyrin
34	305	3.4	606	2 AC2508	hypothetical prote
35	300.5	3.4	1411	2 S30355	alpha-latroinsecto
36	300.5	3.4	2004	2 D88948	protein ZK1005.1 f
37	300	3.4	1058	2 D82654	ankyrin-like prote
38	293.5	3.3	627	2 E86190	hypothetical prote
39	288.5	3.2	662	2 E84725	ankyrin-like prote
40	288	3.2	1184	2 T00253	gene Ankhzn protei
41	286	3.2	517	2 T48283	ankyrin-like prote
42	285	3.2	368	2 T18184	ankyrin repeat pro
43	285	3.2	633	2 T37499	hypothetical prote
44	284	3.2	1083	2 S48460	probable membrane
45	280	3.2	1001	2 S30385	G9a protein - huma

#### ALIGNMENTS

#### RESULT 1

T43458

hypothetical protein DKFZp434F0621.1 - human

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T43458

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z22517

A:Accession: T43458

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1031 <AAA>

A:Cross-references: UNIPROT:Q9UF42; EMBL:AL133620

A:Experimental source: adult testis; clone DKFZp434F0621

C:Genetics:

A>Note: DKFZp434F0621.1

Query Match 58.6%; Score 5205.5; DB 2; Length 1031;

Best Local Similarity 98.5%; Pred. No. 4.9e-284;

Matches 1008; Conservative 2; Mismatches 0; Indels 13; Gaps 1;

QY	1	MSVLISQSVINVEENIPAL	KALLEKCKVDNERECGQTP	LMIAAEOGNLEIVKELIKN	60
Db	1	MSVLISQSVINVEENIPAL	KALLEKCKVDNERECGQTP	LMIAAEOGNLEIVKELIKN	60
QY	61	GANCNLELDNWN	TALISASKEGHVHIV	VEELKCGVNLHHRDMGGWTAL	120
Db	61	GANCNLELDNWN	TALISASKEGHVHIV	VEELKCGVNLHHRDMGGWTAL	120
QY	121	ELLISHGANPSVT	GLQYSVYPIIWAAGRHADI	IVHLLQNGAKVNCSDKYGTTP	180
Db	121	ELLISHGANPSVT	GLQYSVYPIIWAAGRHADI	IVHLLQNGAKVNCSDKYGTTP	180
QY	181	KHLECVKHLAMGAD	VDQEGANSMTALIV	AVKGGYTOSVKEILKRNPNVNL	240
Db	181	KHLECVKHLAMGAD	VDQEGANSMTALIV	AVKGGYTOSVKEILKRNPNVNL	240
QY	241	LMIAASKEGTEI	VQDLLDAGTYVNI	IPDRSGDTVLIGAVRGGHVEIV	300
Db	241	LMIAASKEGTEI	VQDLLDAGTYVNI	IPDRSGDTVLIGAVRGGHVEIV	300
QY	301	QDNKTALYWA	VEKGNATVRDILQCN	PTEICTKDGETPLIKATQNRNIE	360
Db	301	QDNKTALYWA	VEKGNATVRDILQCN	PTEICTKDGETPLIKATQNRNIE	360
QY	361	KVSAVDKGGD	PLHIAIGRSRKLAE	LLRNPKGRLLYRNPKAGETPN	420
Db	361	KVSAVDKGGD	PLHIAIGRSRKLAE	LLRNPKGRLLYRNPKAGETPN	420
QY	421	TOIFGARHLS	PTETDGMGLGYD	LYSSALADILSEPTMOPPIC	480
Db	421	TOIFGARHLS	PTETDGMGLGYD	LYSSALADILSEPTMOPPIC	480



[illegible]

RESULT 3  
T46392  
hypothetical protein DKFZp434E0620.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C:Accession: T46392  
R:Ottewaelder, B.; Obermaier, B.; Meves, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23031  
A:Accession: T46392  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-246 <AAA>  
A:cross-references: UNIPROT:Q9NT37; EMBL:ALJ37553  
A:Experimental source: adult testis; clone DKFZp434E0620  
C:Genetics:  
A>Note: DKFZp434E0620.1

Query Match	14.1%	Score 1255;	DB 2;	Length 246;
Best Local Similarity	99.6%;	Pred. No. 3e-63;		
Matches 246;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1470	SGTEESDNTPLLKDDKDKRKAEGKVERVPKSPHSAEPIRTFIKAKYLSDALLDKDDSSD	1529	
Db				
1	SGTEESDNTPLLKDDKDKRKAEGKVERVPKSPHSAEPIRTFIKAKYLSDALLDKDDSSD	60		
Qy	1530	SGVRSSESSPNHSLHNEVADDSOLEKANLLELEDDSHSGRGIPHSLSGLODPIIARMSI	1589	
Db				
61	SGVRSSESSPNHSLHNEVADDSHLEKANLLELEDDSHSGRGIPHSLSGLODPIIARMSI	1320		
Qy	1590	CSDDKKSPPSCSLIASSPEENWPACQKAYNLNTPSTVTLLNNSAPANRANQFDEMEGI	1649	
Db				
121	CSDDKKSPPSCSLIASSPEENWPACQKAYNLNTPSTVTLLNNSAPANRANQFDEMEGI	180		
Qy	1650	RETSQVILRPSSSPNPTTIQENLKSMTXRSQRSSYTRLSKDPPPELHAAASSESTGFG	1709	
Db				
181	RETSQVILRPSSSPNPTTIQENLKSMTXRSQRSSYTRLSKDPPPELHAAASSESTGFG	240		
Qy	1710	ERESIL 1715		
Db				
241	ERESIL 246			

RESULT 4  
SJHUK  
N;Alternate names: ankyrin 1, erythrocyte splice form 1 - human  
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R

N;Contains: ankyrin 2.2  
C/Species: Homo sapiens (man)  
C/Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text\_change 09-Jul-2004  
C/Accession: A33219; A33219; PC2220; A35443  
R;Lux, S.E.; John, K.M.; Bennett, V.  
Nature 344, 36-42, 1990  
A/Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure w  
A/Reference number: S08275; MUID:90158830; PMID:2137557  
A/Accession: S08275  
A/Molecule type: mRNA  
A/Residues: 1-1881 <LU1>  
A/Cross-references: UNIPROT:P16157; EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702  
A/Accession: A33219  
A/Molecule type: protein  
A/Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753;828-833,'X',835-855,'X'  
x',1367;1383-1427;1601-1630;1686-1698,'D',1700;1763-1772 <LUX>  
A/Note: 845-Arg and 1392-Thr were also found  
R;Hermann, J.; Barel, M.; Frade, R.  
Biochem. Biophys. Res. Commun. 204, 453-460, 1994  
A/Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membrane  
A/Reference number: PC2220; MUID:95071348; PMID:7526850  
A/Accession: PC2220  
A/Molecule type: protein  
A/Residues: 910-929 <HER>  
R;Davis, L.H.; Bennett, V.  
J. Biol. Chem. 265, 10589-10596, 1990  
A/Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchanger  
A/Reference number: A35443; MUID:90285190; PMID:2141335  
A/Accession: A35443  
A/Molecule type: protein  
A/Residues: 'X',5,'X',7-12;403-417,'X',419-422,'H',424,'LQ',797-800,'L',802-814;862-863,  
C/Genetics:  
A/Gene: GDB:ANK1; ANK  
A/Cross-references: GDB:118737; OMTM:182900  
A/Map position: 9p11.2-9p11.2  
C/Supfamily: ankyrin; ankyrin repeat homology  
C/Keywords: alternative splicing; phosphoprotein  
F/2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>  
F/2-1532,1675-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>  
F/2-827/Domain: 89K #status predicted <DOM1>  
F/2-827/Region: anion exchange protein binding  
F/44-76/Domain: ankyrin repeat homology <AN01>  
F/77-109/Domain: ankyrin repeat homology <AN02>  
F/110-142/Domain: ankyrin repeat homology <AN03>  
F/143-171/Domain: ankyrin repeat homology <AN04>  
F/172-204/Domain: ankyrin repeat homology <AN05>  
F/205-237/Domain: ankyrin repeat homology <AN06>  
F/238-270/Domain: ankyrin repeat homology <AN07>  
F/271-303/Domain: ankyrin repeat homology <AN08>  
F/304-336/Domain: ankyrin repeat homology <AN09>  
F/337-369/Domain: ankyrin repeat homology <AN10>  
F/370-402/Domain: ankyrin repeat homology <AN11>  
F/403-435/Domain: ankyrin repeat homology <AN12>  
F/436-468/Domain: ankyrin repeat homology <AN13>  
F/469-501/Domain: ankyrin repeat homology <AN14>  
F/502-534/Domain: ankyrin repeat homology <AN15>  
F/535-567/Domain: ankyrin repeat homology <AN16>  
F/568-600/Domain: ankyrin repeat homology <AN17>  
F/601-633/Domain: ankyrin repeat homology <AN18>  
F/634-666/Domain: ankyrin repeat homology <AN19>  
F/667-699/Domain: ankyrin repeat homology <AN20>  
F/700-732/Domain: ankyrin repeat homology <AN21>  
F/733-765/Domain: ankyrin repeat homology <AN22>  
F/766-798/Domain: ankyrin repeat homology <AN23>  
F/828-1382/Domain: 62K #status predicted <DOM2>  
F/1383-1881/Domain: 55K #status predicted <DOM3>

Query Match 6.3%; Score 555.5; DB 1; Length 1881;  
Best Local Similarity 19.7%; Pred. No. 1.6e-22;  
Matches 374; Conservative 286; Mismatches 659; Indels 581; Gaps 72;  
QY 21 LKALLEKCKDVRNECGQPLMIARQGNLEIWKELIKNGANCNELDNWTALISASK 80

Db 94 VRELNVYGANVNAQSGKFTPLWAAQENHLEVVKFLENGANQONVATEDGFTPLVALQ 153  
QY 81 EGVHVIIEELLKGV-----NLHRDMGCGWTALMMA 111  
Db 154 QGHENVVAHLINYGKGVKRLPALHIAARNDDRTAAVLLQNDPNPDVLSKTGFTPLHIA 213  
QY 112 CYKGRDVELLSHGANSPT-----GLOY-----SV 139  
Db 214 AHYENLVNAQLLNRGASVNFPTQNGITPLHIAASRRGNVIMVRLLLDRGAQIEFTKDEL 273  
QY 140 YPIIWAAGRGHADIHLLQLONGAKVCSKGYTTPLVWAARKHLECVKHLLAGADV- 198  
Db 274 TPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQOHDLCVRLQLQYDAIDD 333  
QY 199 -----OEGA-----NSMTALIVAKVGGYTQSVKILKR 226  
Db 334 ITLDHLTPLHVAACHGHRVAKVLLDKGAKPNRSLNGFTPLHIAACKNHVRVWELLKT 393  
QY 227 NPNVNLTKDQNTALMIASKEGHEIVQDILLDAGTYVNPDRSGDVTILGAVRGHVEIV 286  
Db 394 GASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVETPLHMAARAGHTEVA 453  
QY 287 RALLOKYADIDIRQDKTALYAVEKGNATMYRDILQCNPDTEICTDGETELIKATKM 346  
Db 454 KYLLQNKAKVNAKQDQTPHCAARIGHTNMWKLLENANFNLAITAGHTPLHIAARE 513  
QY 347 RNIEVVELLDKGAKVSAVDKGDTPHIAAIRGSRKLAELLRNPKDGRLLYRPNKAGE 406  
Db 514 GHVETVLALLEKEASQACMTKGFTPLHVAKYGVKVAELLE--RDA---HPNAAKG 567  
QY 407 ---TPNIDCSHOKSILTQIFGARHLSPTETDGMIGY-DLYSSALADILSEFTMQPPIC 462  
Db 568 NGUTPLHVAVHNHNDIVKLLPRGSGSPHSPAWN--GYTPLHTAAQN-----QVEVA 618  
QY 463 VGLYAQWGSKSLKLEDEMTKTFAGQOIEPLFPQSWLIVFLTLKCG-----LGLLFA 518  
Db 619 RSLQYGGSNAAESV-----QGVTPH-----LAAQEGHAEMVALLIS 656  
QY 519 FTVHPNLGIAVSLFIALYIFPIVYFGRREGESNMMAWLSRLARHIGYLELLK- 577  
Db 657 KOANGNLGNKSGLTPLHLV-----AQSG-----HVPVADVLIK 690  
QY 578 -----LMFVNPPPELPGOTTALPVRLF-----TDYNRLSSVG-----GTSLA 616  
Db 691 GVMVDATTRMGYTPLHVASHYGNIKLVKFLLOHQADVNAKTGLGYSPLHQAQOQHTDIV 750  
QY 617 EMI-----ATLSDACERBFGFLA-FRLFRVFKTEDTOGKKWKCTCCLPSFV 662  
Db 751 TLLKNGASPNVSSDGTTPPLATAKRLGYISVTDVLKV-TDET-----SFV 796  
QY 663 IFLFIIGCIISGITLAIAPRVDPKHL-----TVNAVLIASIVVG---LAFVLCRTWMQ 714  
Db 797 L-----VSDKHRMSPPETVDEIL-DVSEDEGELISFKAER- 832  
QY 715 VLDSLINSQKRLHNAASKLHLKSGFMKVLKCEVELMARMAKTTIDSTFQNTRLVVI 774  
Db 833 --DSRDVDEEKELDFVKLDQVVESPAIPRIFC-----AMPETVIRSEEQOASKEY 884  
QY 775 DGLDACEQDKVLQMLDTRVRLFS--KGPFTAIASDP-----HIK- 814  
Db 885 DEDSLPSSPATSTSDNISPVASPVHTGFLVSPWDARGSMRGSRHNGLRVUIPRTCA 944  
QY 815 -----AINQNLNSVLRDSNIGHDYMRN-IVHLPVFL-NSRGLS 851  
Db 945 APTRITCRLVKPKLSTPPPLAEEGLASRIIALGTGAQFLSPVIVEIPHAFSHGKG- 1002  
QY 852 NARKFLVTSATNGDVPCSDTTGQEDADRVSQNSIGEMTKLGSALTALNRDTRRRQW 911  
Db 1003 -DRELVLVRSNGSV-----WKHRSRYGSEYLDQILN-GWDBELGSLLEEKKRVC 1052  
QY 912 RTITRQMSFOLTKLLVT---EDWFSDISPOTMRLLNIVSVTGRLLRANQISFNWDLAS 968

Db 1053 RIIT--TDPLFYFVMSRLQCD-YDIIGPEG-----GSLSKSLVPLVQATPPENAVTK 1102  
QY 969 WINLTQWMPYRTSWLILYLEETEGIPDQMTLK-----TIYERISK----- 1008  
Db 1103 RVKLAQ-----AQPVDELVTLLGNQATFSPIVTEPRRRKPHRPIGLR 1148  
QY 1009 -NIPITTKVEPLLEIDGDIRNFEV-----LSSRTPVLVARDVKVLPCTV 1053  
Db 1149 IPLPFSWTNPRDSGEGDTLSRLLCVIGGTDQAQWEDITGTTLKUYANECANF---TT 1205  
QY 1054 NIDPKLRETIADVRAAREQISIGGLAYPPLPLHGEPPRAPSGYSGSPSSCSSTSFNGPFA 1113  
Db 1206 NV--SARFWLSDCPRTAEAVFATLLYKEL-----TAPVYMAKFV 1243  
QY 1114 GGVSQPHSSYYSGMTGPHQHPFYNGSGPAPGPV--VLLNSLVNDVACE-----K 1162  
Db 1244 -----IFAKMNDPREGLRCYCMTDDKVDKLTQEHENFVEVAR 1281  
QY 1163 LKQIEGLDQSMPLPOYCTTIKKANINGRVLACQNIDELKKEMNMNFGDHLFRSTVLEMRN 1222  
Db 1282 SRDIEVLEMSL-----FAELSG-----NLVPVKAAQQRSPHFOSFRENRLAM-- 1325  
QY 1223 AESHVVPEDPRFLSESSSGPAPHGEPARRASHNE-----LPHTELS----- 1263  
Db 1326 -----PVKVRDSSREPFGGSLFLRKAMKYEDTQHILCHLNTMPPCAKSGAEDRR 1376  
QY 1264 -SOTPYTLNFSPEELNTLGLDEGAPR-----HSNLWSQSQTRRTP-SLSSLSNQD 1311  
Db 1377 RTPTPLALRYSILSESTPGSLSGTEQAEMKMAVISEHLGLSWAELARELQFSVEDIN-- 1433  
QY 1312 SSIEISKLTQVQAEYDAYREVIQMSOL-----EGPGSTTISGRSSPHSTYMGQSSS 1367  
Db 1434 -----RIRVENNSLLSQSVALLNLWIREG-----QNNANENLTALQSID 1475  
QY 1368 GGSIHSLNQEKGKSEPKPDDGRKSFMKRGVDYDYSYSGSVSTNDASPLDPIITEDE-- 1425  
Db 1476 RGEIVNMLESGRQSRNLKPD-----RRHTDRDYSLSPSQMNGYSSL-----QDELL 1522  
QY 1426 -----KSDQSGSKL-----LPGKKSERSSLFQTDLKUGSGGLRYOKPLPSDE 1467  
Db 1523 SPASLGALCALSPURADQYWNVAVLDAIPLAATAEDHTMLESDMQVWSAGLTPLSLVTAED 1582  
QY 1468 DE---SGTESDNTPLKDDKDKKAGKVERVPKSPHSAEPIRTFIKAKYLSDALDK 1524  
Db 1583 SSLECSKAEDSDAT-----GHEWKEGALUSEPRGPE-----LGSLELVEDDTVDS 1628  
QY 1525 KDSS-----DSGVRSSESPNHSNHNEVADDSQLEKANLIELEDDSHSGKRGIHPSL 1576  
Db 1629 DATNGLIDLLEQEGQRSEKPLPGSKRQDDATGAGQ--DSENEVSLVSGHQRGQARITHS- 1686  
QY 1577 SGLQDDPIIARMSICSEDK-----KSPSECSLIASSPEENW 1611  
Db 1687 -----PTVSQVTERSQDRQLQWDADGSIIVSYLQDAAQGSW 1721

## RESULT 5

B35049

ankyrin 1, erythrocyte splice form 3 - human

N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R

N;Contains: ankyrin 2.2, erythrocyte

C;Species: Homo sapiens (man)

C;Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 10-Jul-1998

R;Accession: B35049

R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, M.C.;

Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990

A;Title: cDNA sequence for human erythrocyte ankyrin.

A;Reference number: A35049; MUID:90175370; PMID:1689849

A;Accession: B35049

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1856 &lt;LAM&gt;

C;Genetics:

A;Gene: GDB:ANK1; ANK



QY 1426 -----KSDQSGSKL-----LPGKSSRSLSLFTQDLKLKSGSLRYQKLPSDE 1467  
Db 1523 SPASLGALSSPRAQYQWNEVAILDAIPLAATHEHDTMLEMSDMQWWSAGLTPLSLVTAED 1582  
QY 1468 DE---SCTRESNTPLLKODKDKABGKVERVPKSPHSAEPIRTFIKAKEYLSLALLDK 1524  
Db 1583 SSLESCAEDSDAT-----CHEWLEGALSEERPRPE-----LGSILEVEDDTVDS 1628  
QY 1525 KDS-----DSGVRSESSNNHLSHNEVADDSQLEKANLIEBDDSHSGKRGIPHS 1576  
Db 1629 DATNGLIDLEQEGORSEELPGSKRQDDATGAGQ--DSENEVSLVSGHORGQARITHS-- 1686  
QY 1577 SGLQDPIIARMSICSEDK-----KSPSECSLIASSPEENW 1611  
Db 1687 -----PVSQVTSRQDRQDWDAGSIVSYLQDAAGGSW 1721

## RESULT 6

A35049  
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R  
N;Contains: ankyrin 2.2, erythrocyte  
C;Species: Homo sapiens (man)  
C;Date: 27-Jul-1990 #sequence\_revision 01-Oct-1992 #text\_change 09-Jul-2004  
C;Accession: A35049  
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, M.C.; K  
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990  
A;Title: cDNA sequence for human erythrocyte ankyrin.  
A;Reference number: A35049; MUID:90175370; PMID:1689849  
A;Accession: A35049  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1880 <LAMB>  
A;Cross-references: UNIPROT:P16157; GB:M28880  
C;Genetics:  
A;Gene: GDB:ANK1; ANK  
A;Cross-references: GDB:118737; OMIM:182900  
A;Map position: 8p11.2-8p11.2  
C;Superfamily: ankyrin; ankyrin repeat homology  
C;Keywords: alternative splicing; cytoskeleton  
F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>  
F;2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>  
F;44-76/Domain: ankyrin repeat homology <AN01>  
F;77-109/Domain: ankyrin repeat homology <AN02>  
F;110-142/Domain: ankyrin repeat homology <AN03>  
F;143-171/Domain: ankyrin repeat homology <AN04>  
F;172-204/Domain: ankyrin repeat homology <AN05>  
F;205-237/Domain: ankyrin repeat homology <AN06>  
F;238-270/Domain: ankyrin repeat homology <AN07>  
F;271-303/Domain: ankyrin repeat homology <AN08>  
F;304-336/Domain: ankyrin repeat homology <AN09>  
F;337-369/Domain: ankyrin repeat homology <AN10>  
F;370-402/Domain: ankyrin repeat homology <AN11>  
F;403-435/Domain: ankyrin repeat homology <AN12>  
F;436-468/Domain: ankyrin repeat homology <AN13>  
F;469-501/Domain: ankyrin repeat homology <AN14>  
F;502-534/Domain: ankyrin repeat homology <AN15>  
F;535-567/Domain: ankyrin repeat homology <AN16>  
F;568-600/Domain: ankyrin repeat homology <AN17>  
F;601-633/Domain: ankyrin repeat homology <AN18>  
F;634-666/Domain: ankyrin repeat homology <AN19>  
F;667-699/Domain: ankyrin repeat homology <AN20>  
F;700-732/Domain: ankyrin repeat homology <AN21>  
F;733-765/Domain: ankyrin repeat homology <AN22>  
F;766-798/Domain: ankyrin repeat homology <AN23>

Query Match 6.2%; Score 552.5; DB 2; Length 1880;

Best Local Similarity 19.6%; Pred. No. 2.4e-22;  
Matches 373; Conservative 287; Mismatches 659; Indels 581; Gaps 72;

QY 21 LKALBKCKDVERNECGQTPLMAAEQGNLEIVKELIKNGANCNLELDNNTALISASK 80  
Db 94 VRELNVGANVNAQSKGFTPLYMAAQENHLEVVKFLENGANQNVATEDGFTPLAVALQ 153

QY 81 EGHVHIVEELLKGV-----NLEHRDMCGWTALMWA 111  
Db 154 QGHENVAHLINYGTKGVRLPALHIAARNDDTRTAAVLLQNDPNDVLSKGTGFTPLHIA 213  
QY 112 CVKGRTDVVVELLSHGANSVT-----GLQY-----SV 139  
Db 214 AHYENLVAQLLNRGSSVNFPPQNGITPLHIASRRGNVIMVRLLLDRGAQIETKTDEL 273  
QY 140 YPIIWAAGHADIIVHLLONGAKVNCSDKYGTGTPLVWAARKGHLECVKHLLAMGADV- 198  
Db 274 TPLHCAARNGHVRISBILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVCRVLLQYDAEIDD 333  
QY 199 -----QEGA-----NSMTALIVAKGVYQTSVKELIKR 226  
Db 334 ITLDHLTPLHVAACHGHHRVAKVLLDKGAKPSRALNGFTPLHIACKNHRVREULLKT 393  
QY 227 NPNVNLTDKGNLTALMIASKEGHEITVODLLDAGTYVNIIPDRSGDTVLICAVRGHVEIV 286  
Db 394 GASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAABAGTEVA 453  
QY 287 RALLOKYADIDIRGQNKATLYAWVEKNATWVRDILQCNPDTEICTKOGETPLIATKM 346  
Db 454 KYLLQNKAKVNAKADQDTPLHCAARIGHTNMVKLLENNANPNLATTAAGHTPLHIAARE 513  
QY 347 RNIEVVELLDKGAKVSAVDKKGDTPLHIAIRCSRKLAELLLRNPKDGLLYRPNKAGE 406  
Db 514 GHVETVLLALLEKEASQACMTKKGFTPLHVAAKYKVRVAELLES--RDA-----HPNAAGK 567  
QY 407 ---TPVNI DCSHQKSI LTOIFGARHLSPTETDGDMLGY--DLYSSALADILSEPTMOPPIC 462  
Db 568 NGLTPLHVAHNNLDIVKLLPRGSGSPHSPAWN--GYTPLHIAAKQN-----QVEVA 618  
QY 463 VGLYAQWGGKGFLLKKLEDEMKTFAQOIPELPQFSWLIIVFLTLILLCGG-----LGILFA 518  
Db 619 RSLLOYGGSANAESV-----QGVTPHL-----LAAQEGHAEMVALLLS 656  
QY 519 FTVHNHGIATVSLFLALYIFVIVYFGRREGESNNWAVLSTRLAHIGVLELLK- 577  
Db 657 KQANGNLNKGSLGTLPLHLV-----AQEG-----HVPVADVLIIKH 690  
QY 578 -----LMFVNPPELPBQTTKALPVRF----TDYNRLSSVG-----GETSLA 616  
Db 691 GUMVDATTENGTYPLHVASHYGNIKLVKLLQHQADVNAKTGLYSPLHQAQOQHTDIV 750  
QY 617 EMI-----ATLSDACEREFGLA--TRLFRVPKTEDTQKKWKKTCCCLPSFV 662  
Db 751 TLLLNKGASPNESVSDGTTPLAIAKRLGYISVTDVLKVV--TDET-----SFV 796  
QY 663 IFLFIIGCIISGITLAIIPRVDPKHL-----TVNAVLISIASVVG---LAFVLNCRTWQ 714  
Db 797 L-----VSDKHMSPPETVDEIL--DVSEDEGELISFKAER- 832  
QY 715 VLDSLINSQKRLHNAASKLHLKSEGFMKVLKCEVELMARMAKTIDTSQNTQRLVVI 774  
Db 833 --DSRDVDEKELLDFVPLKLDQVVESPAIPRPC-----AMPETVIRSEEQEASKEY 884  
QY 775 DGLDACEQDKVQLMDTVRVLFS--KGPFIATPASDP-----HIILK-----814  
Db 885 DEDSLIPSSPATETSDNISPVASPVHTGFLVSPVDARGSMGRSHRNLGRVVPRTCA 944  
QY 815 -----AINQNLNSVLRDSNINGHDYMN--IVHLPVEL--NSRGLS 851  
Db 945 APTRTICRLVPOKQSTPPPLAEEBGLASRIIALGPTGAQFLSPVIVEIPHFAHSGRG-- 1002  
QY 852 NARKFLVTSATNGDVPSCSDTTGQEDADRRVSONSLGEMTKLGSKTALNRDRYRRQM 911  
Db 1003 -DRELVLASENGSV-----WKEHRSRYGESYLDQILN--GMDEELGSLSELEKRV 1052  
QY 912 RTITRQMSFDLTKLIVT---EDWFSDISQTMRRLLINIVSVTGRLLRANQISVNRDLAS 968  
Db 1053 RIIT--TDPLPVFVIMSRLCQD--YDIIGPEG-----GSLSKSLVPLVQATPENAVTK 1102







Db 1449 TETSVL---KSH-----LVNEVPVLASPDLLSEVSEMKGQDLIKM---TAILT 1489  
Qy 1184 ANINGRVLACQNMIDELKEMNMNFGDMHLFRSTVLEMRNAESHVVPDPFLSE--SSGCP 1242  
Db 1490 TVSDKA--GSIKVELVKAEEEPGE-----PPEIVER-----VKEDLEKVNELRSQT 1537  
Qy 1243 APHGEPARRASHNELPHTELSSQTPYTLNFSEELNLTGLDEGAPRHSNLWSQSTRTP 1302  
Db 1538 CTRDESSVQSSRSRGLVE--EEMVIVSDEIEEAR-----QKAPLEITEYPCVEVRIDK 1590  
Qy 1303 SLSSINSODSIEIETSKLTDKQVAEYRDAYRIYAQMSQLEGPGSTTISGRSSPHSTYIM 1362  
Db 1591 EIKGKVEKDTGLVNYLTDLLNT--CVPLPKQLQTVQDKAGKCEALAVGRSS----- 1642  
Qy 1363 QSSSSGGSIHNLQEKKGQSEP-----KPDGCKSFLMKRGDVIDYSSSGVSTN 1412  
Db 1643 -----EKEGKQIPPETQSTQKHKPSLGIKKPVRRK----- 1674  
Qy 1413 DASPLDPITEDEKSDQSGSKLLPKQKSSRSLSFQTDCLKGSGRLRYOKLPSEDESQT 1472  
Db 1675 ---LKEKQKQKEGLQASAEKELKGSSESLGE-----DPLAPEPLPTVKATS-- 1722  
Qy 1473 EESNTPLDKDDKQKABKVERVPKSP-----HSABPIPTFIKAKYELSDALLDKOS 1527  
Db 1723 -----PLIEETPIGSIKDKVKALQKRVDEBQKGRSKLPFR--VKGKEDVPKKTTHRPH 1774  
Qy 1528 SDGVRSE-----SSPNHSLHNEVADDSQLEKANLELEDSDSHGKRGIPHSLGLQD 1581  
Db 1775 AASFSLKSERHAPGSPKTERHSTLSSAKTERHPVS-----PSSKTEKHS 1822  
Qy 1582 PIARMSICSDKSPSCSLIASSPEENWPACQKAYNLNRTPTVTTLNNSAPANRANQ 1641  
Db 1823 PV--SPSAKTERHSPASS--SKTEKHSV-----SPSTKTERHSPVSTKTER 1867  
Qy 1642 NFDEMEGIRETSQVILRPSSPNPTTQENLK--SMTHKRSQRSYRFLSKDPP 1694  
Db 1868 H----PPVSPSGTKDKRPVSPSGRTEKHPPVSPGRTEKRLPVPSPGRTRKHQP 1917  
RESULT 9  
T42716  
ankyrin 3, splice form 4 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T42716  
E:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialanias, M.; Turtzo, I.  
J. Cell Biol. 130, 313-330, 1995  
A>Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene  
the repeat domain.  
A:Reference number: 222237; PMID:95340633; PMID:7615634  
A:Accession: T42716  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1961 <PET>  
A:Cross-references: UNIPROT:Q61307; EMBL:U40632; NID:g710548; PID:g710552; PIDN:AAB01607  
A:Experimental source: strain C57BL/6J, kidney  
C:Genetics:  
A:Gene: Ank3  
A:Map position: 10  
C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing  
Query Match 6.0%; Score 530; DB 2; Length 1961;  
Best Local Similarity 19.9%; Pred. No. 4.7e-21;  
Matches 432; Conservative 279; Mismatches 673; Indels 784; Gaps 90;  
Qy 14 BEENIPALKALKERCKVDNERNECGOTPLTAABQGNLEIVKELIKKGNCLNLDLQNT 73  
Db 66 KEGHVEVVELLQREANVDAAKGNLTAHLAGQAEVVKLVLTNGANVNAQSQGFT 125  
Qy 74 ALISASKEGHVHIYVEELLKCGNLEHDMGWTALMMA CYKGRDTPVVELLI----- 124  
Db 126 PLYMAOENHLEVVRFLDNGASQSLATEDGFTPLAVALQGHQDVVSLLENLTKGVR 185

Qy 125 -----SHGANPSVTGL----- 135  
Db 186 LPALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNINIVATLLNRAAAMD 245  
Qy 136 ---QYSVVPITWAAGRHADIVHLLONGAKVNSDKYGTTPPLVWAARKG----- 182  
Db 246 FTARNDTPLHVASKRGVANNVKLLDRGAKIDAKTRDGLTPLHCGASQHEQVVMELLD 305  
Qy 183 -----HLECVKHLAMGADVQEGANSMTALIVAKGGYQTS 219  
Db 306 RSAPILSKTNKLSPLHMTAQGDHLCVQLLQHNVPDVTNDYLTALHVAACHGCHYKV 365  
Qy 220 VKELIKENPNVLTDKDGNLTALMIASKE----- 247  
Db 366 AKVLLDKKASNAKALNGFTPLHIAKKNRIRVMELLKKGASTQAVTESGLTPIHVAAF 425  
Qy 248 -GHTIEIVQDLDLADAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRGQDNKTA 306  
Db 426 MGHVNIIVSQLMHKGASPNVTNVRGETALHMAARSGAEVVRYLVQDGAQVEAKAKDQTP 485  
Qy 307 LYNAVEKGNATWVRDIIQCNPDTEI CTYKGETPLIKATKMRNIEVVELLLDKGAKVSAD 366  
Db 486 LHISARLGKADIVQOQLQOGASPNAAATTSYTPPLHAAAREGHEDVAFLDGHGASLSITT 545  
Qy 367 KKGDTPLHIAIRGRSRKLAELLRL-----NPKDGRLL----- 398  
Db 546 KKGFTPLHVAKYKLEVASLLLOKASPDNAKSGLTPLHVAAHYNQKVALLLDQGA 605  
Qy 399 --YRPNKAGETPNIDC--SHQKSILTQI--FGAR-----HLSTPTETDGMILGY 441  
Db 606 SPHAAAKNGYTPHLIAAKNQMDIATSLLEYGADANAVTRQGIASVHLAAQEGHVDWVSL 665  
Qy 442 DLYSSALADILSEPTMOP-----PICVGLYAO 469  
Db 666 LLSRANVNLNSKSGLTPLHAAQEDRVNAEVLVNOGAHVDAQTKMGYTPHLVGHCH--Y 723  
Qy 470 GSGK--SFLK--KLEDEMKT-----FAGQQIEPLFPQFMSWLIVFLTLCCGG--LG 514  
Db 724 GNIVNPLQHSKAVNAKNGYTLHQAQAQ-----GHTHIIN 763  
Qy 515 LLFAFTVHPN-----LGIAVLSFLALLYIFFIVI----- 544  
Db 764 VLLQNNASPNELTVNGNTALAIARRLGIVISVVDTLKVVTBEIMTTTTITEKHKNVPEM 823  
Qy 545 -----YFGRREGESNNWAVLSTRLARHIGVLELLKL-- 578  
Db 824 NEVLDMSDDEVKASAPEKLSDEYISDGEEDKCTWFKIPKQV-----EVLVRSKD 875  
Qy 579 -----MFVNPPPELPEQTTKALP-----VRFLETD---YNRLSSVCGE 612  
Db 876 AITGDTKYLGPQDLKELGDDSLPAGYVGFSLGARSASLSRFSRSDSYTLNRSYARDS 935  
Qy 613 TSAEMTATLSDACEREFGLFATLRFVFKTEDTQGGKKKKKTCCLPSFVIFLFIIGCI 672  
Db 936 MMIEELLVP-----SKEQHLTFTEF-----DSDSLRHSWAADT-----LDNVNLS 978  
Qy 673 SGI--TLTALFRVDPKHLTVNAVLIISASVGLAFVL---NCRTWQVLDLSLNSQRKL 727  
Db 979 SPVHSGFLVSMVDARGSMRG-----SRHHGMRIIPPRKCTAPTTRITCLV--KRHL 1031  
Qy 728 HN-----AASKLHKLKSEG-----FMVKLCEVELMARMAKTIDSTQ 765  
Db 1032 ANPPPMVEGSLARVEMGPAGAFGLGPIVIEIPHFGSMRGKREILVISENGETWKE 1091  
Qy 766 NQ-----TRLVIIDGLDACEQDKVLQMLDTPVR--LFSKGPFTAIASDPIIHKAINQ 818  
Db 1092 HQPDKNEDLAELNGND--EELDSPELGTGKRICRIITK--DFQYFA-----VVSRIKQ 1143  
Qy 819 NLNSVLRSNNGHDMYRNIVHLFVFLNSRGLSNARKFLV--TSATNGDVPSCDPTTGQ-- 875  
Db 1144 ESNQIGPEGGI-----LSSTVTPVLQASFPFEGALTKRIRVGLQAO 1183

QY 876 ---EDADRRVQNSLGEWTKLGSKTALNRDVTY--RRQWORTITROMSFDTLTKLL---V 927  
Db 1184 VPPEETVKKI-----LGNKATSPVIVVEPRRRKPKHPIIT--MTIPVPPSGEGV 1231  
QY 928 TEDWFSDISPQTRRLINIVSVTRGLLRANQISFNWDRLASWINLQWPF-----Y 978  
Db 1232 SNGYKGDATP-NLRLLCSITGTSF-----AQWEDICTPTLTIFKDCVSF 1276  
QY 979 RTS-----WL-----ILYLETSGIPDQWMLTKI-YERISKNTPTTKDVEPLEIDGDIR 1027  
Db 1277 TTNVSARFNLADCHQVL---ETVGLASQLYRELICVPYMAKFFVFAKTNDP---VESSLR 1330  
QY 1028 -----NFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLR---EIIADVR 1067  
Db 1331 CFCMTDRVDKTLQEQENFEVEARSK-DIEVLEGKPIYDVCYGNLAPLTKGGQQLVFNFY 1389  
QY 1068 AAREQISIGLAVPLPLHEGPPRPSGY-----SOPPSVCSSTSFNGPPAGGWSFQPH 1122  
Db 1390 SPKEN-----RLPFSIKIRDTSQEP--CGRLSP-----1415  
QY 1123 SSYYSGMTGQHPFYNRGSPAPGPVVLLNSLNVDAVCEKLQIEGLDQSMPLPQYCTTIK 1182  
Db 1416 -----LKEPKTKGLPQTAV-----1430  
QY 1183 KANINGRVLAQCNIDELKKEMNNFGDWHLFRSTVLEMRNAESHVVPEDRFLSESSESGP 1242  
Db 1431 -CNLNTLPLAHPKAERKQRS-----FASLALRKRY-----YLTEPSMSP 1471  
QY 1243 APHGEFARR-----ASHNELPHTLSSQTPYTLNFSPEELNTLGLDEGAPRHNSL- 1293  
Db 1472 Q---SPCERTDIRMAIVADHLGSLWTELARE---LNFVSDEINQIIVEN---PNLSISQ 1521  
QY 1294 -----W---QSOTRTPSLSLNSGDSISIEISKLTDKVOAEB-----RDARYEYIAQ 1337  
Db 1522 SFMLLKQWTRDCKNATTDALTSVLTAKINRIDIIVTLLEGPIFDYGNISGRSPADENNVE 1581  
QY 1338 MSQLEGGPG-----STTISGRSSPHSTYYMGQSSSGSISHS--N 1374  
Db 1582 HDVVDGHPSPQVLETFMGLYWTFPPNPFQDDHFSDISSTESPFRTSPRLSDGLVPSQGN 1641  
QY 1375 LEQEGK-----KDS---EPKPDGKRKFLMKRGDVIDYSSSGVSTNDASPLDPTTEED 1424  
Db 1642 IEHPTGGPPVTAEDTSLEDSKWD-----SVTVDPADPLD--VDES 1682  
QY 1425 EKSQDSGSKL-----LPQKKSSERS--LPQTDLKLKGLGYQKLPDSEDESGTE--- 1473  
Db 1683 QLKDLQCECAQCWASVPGIPNDGRQAEPLRPQTRKVGMSSEQQEKGKSGPDDEVEDKV 1742  
QY 1474 -----ESDNTPLLKDKRKAEGKVERVP-----KSPEHSAB-PIRTFIK 1512  
Db 1743 KSLFEDIQLEVEAEEMTEDQOGAMLNRVQRAELAWSLAWQNETPSPGSLSPAQA--- 1799  
QY 1513 AKEYLSDALLDKDSDSGVRSS-----ESSPNHSLHNEVADDSQLEKANLIE 1560  
Db 1800 --RRLTGLLDRLDDSDQARDISITSLTGEPEKIEANGNHTA--EVIPEA---KAKPYF 1852  
QY 1561 LEDDSHSGKRGK-----PHSLSGIQDPTIARMSICSEDKSPSECS--LIASSPEENW 1611  
Db 1853 PESQNDIGKQS IKENLKPETHGCGRTEEPV---SPLTAYQKSLEETSCLVIEDAPKCV 1908  
QY 1612 PACQKAYNLNRPTSTVLLNNNSAPANRANQFDEMEGIRTSQVILRPSSSPNPTTIONE 1671  
Db 1909 FVGWK--RWTRITAD-----GKARLNLOEEEG-----STRSEPK--QGE 1943  
QY 1672 NLKSMTHK 1679  
Db 1944 GYKVKTK 1951

RESULT 10

T42715

ankyrin 3, splice form 3 - mouse

C,Species: Mus musculus (house mouse)

C,Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C,Accession: T42715  
R,Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, L.  
J. Cell Biol. 130, 313-330, 1995  
A,Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene  
the repeat domain.  
A,Reference number: Z22237; MUID:95340633; PMID:7615634  
A,Accession: T42715  
A,Status: preliminary; translated from GB/EMBL/DBJ  
A,Molecule type: mRNA  
A,Residues: 1-1940 <PET>  
A,Cross-references: UNIPROT:O61307; EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604  
A,Experimental source: strain C57BL/6J; kidney  
C,Genetics:  
A,Gene: Ank3  
A,Map position: 10  
A,Introns: 834/1  
C,Superfamily: ankyrin; ankyrin repeat homology  
C,Keywords: alternative splicing

Query Match 6.0%; Score 529.5; DB 2; Length 1940;  
Best Local Similarity 20.1%; Pred. No. 4.9e-21;  
Matches 431; Conservative 280; Mismatches 673; Indels 763; Gaps 90;

QY 14 ESENIPALKALLECKVDNERNECQOTPLMIAAEOGNLEIVKELIKNGANCNLEDLDNWT 73  
Db 66 KEGHEVVSSELLQREANVDAATKKGNTHLHSLAGAEVVKVLTNGANVNAQSONGFT 125  
QY 74 ALISASKEGHVHVEELLKCGVNLHRDMGGWTALMWACVKGRTDVVELLL 124  
Db 126 PLYMAQENHLEVRFLONGASQSLATEDGFTPLAVALQOQHDDQVVSLLLENDTKGVR 185  
QY 125 -----SHGANPSVTGL----- 135  
Db 186 LPALHIAARKDQTKAAALLQNDTNADVESKSGFTPLHIAAHYGNINVTALLNRAAAYD 245  
QY 136 ---QISVYPIIWAAGRGHADIHLLALONGAKVNCSDKYCTTPLVWAARG- 182  
Db 246 FTARNDDITPLHVASKRGANVMKLLDRGAKIDAKTRDGLTPLHCGARSHEQVEMLLD 305  
QY 183 -----HLECVKHLAMGADVQDEGANSMTALIVAVKGYTQS 219  
Db 306 RSAPILSKTKNGSLPLHMATQGDHLNCVQLLQHNVPDVTNDYLTALHVAACHGHYKV 365  
QY 220 VKEILKRNPNVNLTKDQNTALMIASKE----- 247  
Db 366 AKVLLDKKASPNKALNGFTPLHIAACKNRIRVMELLKHGASIQAVTESGLTPIHVAAF 425  
QY 248 -GHTEIVQDLDAGTYVNIPIRDSRGDTVLIGAVRGHVEIVRALLQKVADIDIRQDNKTA 306  
Db 426 MGHVNIVSQLMHHGASPNNTNVRGETALHMAARSGQAEVVRVLYVQDGAQVEAKAKDDQTP 485  
QY 307 LYWAVEKGNATVMVRDILQCNPDTEICTKDGETPLIKATKRNIEVVELLLDKGAKVSAVD 366  
Db 486 LHISARLGKADIVQQLQOQASPNATTSYTPHLHAAEGHEDVAFLDDHGCASLITT 545  
QY 367 KGGDTPHLHIAIGRSKLAELLR-----NPKDGRLL----- 398  
Db 546 KKGFTPLHVAAYKYGLEVASLLIQKSASPDAAKSGLTPLHVAAHYDQNKVALLDDQGA 605  
QY 399 --YRPNKAGETPVNTDC-SHOKSILTQI--FGAR-----HLSPTETDGMGLY 441  
Db 606 SPHAAAKNGYTPHLHIAAKKNQMDIATSLLEYGADANAVTROGSIASVHLAAQEGHVDNVL 665  
QY 442 DLYSSALADILSEPTMQP-----PICVGLYAOV 469  
Db 666 LLSRNANVNLNKSGLTPLHAAQEDRVNVAEVLNVNQAQAHVDAQTKWGYTFPLHVGHCH--Y 723  
QY 470 GSGK--SFLKK---KLEDEMT-----FAGQOIEPLFOFSWLIIVFTLLLCGG-----LG 514  
Db 724 GNIKIVNFFLQHSKAVNAKTNGYTAHQAAQO-----GHTHIIN 763  
QY 515 LLPAFTVPHN-----LGIAVSLSLFALLIYPIFIV----- 543

Db 764 VLLQNNASPNELTVNGNTALAIARRGLYISVVDTLKVVTEETMTTITTEKHQNVPEM 823  
Qy 544 --IYFGGRREGESMNAWVLSRLARHIGYLELLKL-----MFVNPELPQOTTK 592  
Db 824 NEVLDMSDEGDKCTWFKPKVQ-----EVLKSEDAITGTDKYLGFQDLKELGDD 875  
Qy 593 ALP-----VRLFTD-----YNRLSSVGGETSLAEMIAIATSLDACEFGL 633  
Db 876 SLPAEGYVGFSLGARSASLRSFSDRSYTLNRSSYARDSMMIEELLVP-----SKEQHLT 930  
Qy 634 ATRLFRVPEKTEDTQKKKKWKTCCCLPSFVFLFIIGCIISGI--TLIAIFRVPDPKHLTV 691  
Db 931 FTREF-----DSDSLRHYSAADT-----LDNVLVSSPVGSHGFLVFMVDARGGSMR 978  
Qy 692 AVLISIASVGGLAFVL--NCKTWQVLDLSLNSORKEHLN-----AASKLHLKLS 739  
Db 979 G-----SRHGWRIIIPRKCTAPTRICRLV--KRHLKLANPPWVEGEGSLARLVEMGP 1031  
Qy 740 EG-----FMKVLKCEVELMARMARKTIDSTFQNG-----TRLVVIIDGLDACE 781  
Db 1032 AGAQLGFPVIVEIPHFGSMRGKRELIVLRSENGETKEHQFDSKNEDLAELLNGMD--E 1089  
Qy 782 QDKVLQMLDTRV--LPSKGFPIAFSDPHIILKAINQNLNSVLURDSNNGHDMYMRNV 839  
Db 1090 ELDSPEELGTKRICRIITK-DPPQYFA-----VWSRIKQESNQIGPEGGI-----1133  
Qy 840 HLPVPLNSRGLSNARKFLV--TSATNGDVPDCSDTTGQ-----EDADRVSQNSLGEWTKL 893  
Db 1134 -----LSSTTVPVQASFPEGALTKRIRVGLQAQVPVEETVKKI-----L 1173  
Qy 894 GSKTALNRDITY--RRRQMQRTIROMSFDITKLL---VTBDFSDISPMRRLNIVS 948  
Db 1174 GNKATFPIVTEPRRRKFHPIT--WTIPVPPSGEGVSGYKGDATP-NRLLLCSITG 1230  
Qy 949 VTGRLLRANQISFNWRLASMINITEOMP-----YRTS-----WL-----ILYLEE 989  
Db 1231 GTSP-----AQWEDITGTPLTFIKDCVSTFTNVSGARFWLADCHQVL--E 1273  
Qy 990 TEGIPDQMLAKTI-YERISKNIPTTKOVEPILLEIDGDIR-----NFEV 1031  
Db 1274 TVGLASQYRELICVYNNAKFWVFAKTNDP---VESSLRCFCMTDDRDVDTKLEQOENPEE 1330  
Qy 1032 FLSSRTPLVARDVKVFLPCTVNIIDPKLR---EIIADVRAAREQISIGGLAYPPLPLEHG 1088  
Db 1331 VARSK-DIEVLEGREPIYDCYGNLAPLTGGQQLVFNFSYFKEN-----1373  
Qy 1089 PPRAPSGY-----SQPPSCVSTSGFNGFPAGGVSPQPHSYSGMTGPQHPFYNRGSGP 1143  
Db 1374 --RLPFSIKIRDTSOEP--CGRLSF-----1394  
Qy 1144 AGPFWLLNSLVDAVCEKLOIBGLDQMLPOYCTTIKKANINGRVLACQNBELKEM 1203  
Db 1395 -----LKEPRTTKGLPQTAV-----CNLTITLPAHKKAEDARRQ 1429  
Qy 1204 NMFGDWHLPSTVLEMRNAESHVVPEDRFLSSBSSGAPHPGPARP-----ASHN 1255  
Db 1430 S-----FASLARKYS-----VLTPFSMSPO--SPCETDTRMAIVADHL 1468  
Qy 1256 ELPHTLESSTQPTYNLNFSEELNTGLDEGAPRHSNLS-----W-----OSQTRRTPSL 1304  
Db 1469 GLSWTELARE---LNFSDVEINQIRVEN---PNSLISQSFMLLKKWVTRDGNATTDAL 1521  
Qy 1305 SSLNSQSSSIISIKLTKQAEY-----RAYREYIAQMQLSGGPG-----1346  
Db 1522 TSVLTKNIRIDIVTLLEGPIFDYGNISGTRSFADENNVPDVGHPFSQVELETPMGLY 1581  
Qy 1347 -----STTISRSSPHETTYMGQSSSGSITHS--NLQKRG-----KDS---EP 1385  
Db 1582 WTPNPFQDDHFDSDISSIESPFRTPSKLSGLVPSQGNIEHPTGGPVVTAEDTSLDS 1641  
Qy 1386 KPDDGRKSFLMKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKL-----LPCKK 1439

Db 1642 KMDD-----SVTTDPADPLD--VDESQKDLQCSQAQWASVPGIP 1682  
Qy 1440 SSERSS--LFQDLKLGSLRYQKLPSEDESGETE-----ESDNTPLLKDDK 1485  
Db 1683 NDGEQAPLRLPQTRKVGMSSEQQKSGPDEEVTEKVKSLFEDIQLEEVAEEMTDDQ 1742  
Qy 1486 DRKAEGKVERVP-----KSPHSAB-PIRTFIKAKKEYLSDDLKKKSSDSGVR 1533  
Db 1743 QQAMLNVRQARLAMSLSLAGWQNETPSSGLESQAQ-----RRLTGGLLRLDDSSDQAR 1797  
Qy 1534 SS-----ESSPNHSLNEVADDSOLEKANLIELEDDSHSGKRG-----PH 1574  
Db 1798 DSITSYLTGEGKTEANGNHTA--EVIPEA---KAKPYFPESQNDIGKQIKENLKPKTH 1852  
Qy 1575 SLSLGQPIIARMSICEDKKSPSECS--LIASSPEENWPAQCAKAYNLNRTPTSTVTLNN 1632  
Db 1853 GCGRTEFPV-----SPLTAYQKSLSETSKLVIEDAPKCPVPVGMK--KMTRTTAD-----1900  
Qy 1633 SAPANRANQNFDEMEGIRETSQVILRPSSESNPTTIQNIENLKSMTHK 1679  
Db 1901 ----GKARLNQEBEG-----STRSEPK--QGEQYKVKTKK 1930

RESULT 11  
149502  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I49502  
R:White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.  
Mamm. Genome 3, 281-285, 1992  
A:Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory domain  
A:Reference number: I49502; MUID:92345717; PMID:1386265  
A:Accession: I49502  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1862 <RES>  
C:Cross-references: UNIPROT:Q02357; GB:M84756; NID:g191939; PID:AAA37236.1; PID:g191940  
C:Genetics:  
A:Gene: Ank-1  
C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing  
F:40-72/Domain: ankyrin repeat homology <AN01>  
F:73-105/Domain: ankyrin repeat homology <AN02>  
F:106-138/Domain: ankyrin repeat homology <AN03>  
F:139-167/Domain: ankyrin repeat homology <AN04>  
F:168-200/Domain: ankyrin repeat homology <AN05>  
F:201-233/Domain: ankyrin repeat homology <AN06>  
F:234-266/Domain: ankyrin repeat homology <AN07>  
F:267-299/Domain: ankyrin repeat homology <AN08>  
F:300-332/Domain: ankyrin repeat homology <AN09>  
F:333-365/Domain: ankyrin repeat homology <AN10>  
F:366-398/Domain: ankyrin repeat homology <AN11>  
F:399-431/Domain: ankyrin repeat homology <AN12>  
F:432-464/Domain: ankyrin repeat homology <AN13>  
F:465-497/Domain: ankyrin repeat homology <AN14>  
F:498-530/Domain: ankyrin repeat homology <AN15>  
F:531-563/Domain: ankyrin repeat homology <AN16>  
F:564-596/Domain: ankyrin repeat homology <AN17>  
F:597-629/Domain: ankyrin repeat homology <AN18>  
F:630-662/Domain: ankyrin repeat homology <AN19>  
F:663-695/Domain: ankyrin repeat homology <AN20>  
F:696-728/Domain: ankyrin repeat homology <AN21>  
F:729-761/Domain: ankyrin repeat homology <AN22>  
F:762-794/Domain: ankyrin repeat homology <AN23>

Query Match 5.9%; Score 525; DB 2; Length 1862;  
Best Local Similarity 20.1%; Pred. No. 8.2e-21;  
Matches 396; Conservative 272; Mismatches 648; Indels 654; Gaps 77;  
Qy 21 LKALLEKCKDVERNECGQTPMLTAAECGNLEIVKELIKGANCNLELDNWTALISASK 80  
Db 90 VRELNVNGANVNAQSKGFTPLYMAAQENHLEVVKFLENGANQNVATEDGFTPLAVALQ 149

QY 81 EGHVHIVEELLKGV-----NLEHRDMGCTALMWA 111  
Db 150 QGHENVAHLINLYGTGKGVLPALHIAARNDDTRTAVALLQNDPNPDVLSKTGFTPLHTA 209  
QY 112 CYKGRDVBVLLSHGANPSVT-----GLOV-----SV 139  
Db 210 AHYENLVAQLLNLRGASVNFTPQNGITPLHIAARRGNVIMVRLLLDRGAQIETRYKDEL 269  
QY 140 YPIIAAGRHADIVHLLQLONGAKVNSDKYGTTPPLVMAARKGHLECVHLLAMGADV- 198  
Db 270 TPLHCAARNGHVRISIELLDHGAPIOAKTKNGLSPIHMAAQGDHLLDCVRLLLQYNAEIDD 329  
QY 199 -----QEGA-----NSMTALIVAKGGYTSQVKELKR 226  
Db 330 ITLDHITPLHVAACHGHRHVRKVLLDKGAKNRSLNGFTPLHIAACKNHIRVWELLKKT 369  
QY 227 NPVNLTDKDGNLTALMASKEGHEITVQDLDLAGTYVNIPIDRSGDTVLIGAVRGHVEIV 286  
Db 390 GASIDAVTESGLTPLHVASPMGHLPIVKNLLQRGASPNVSNKVEITPLHMAARAGTEVA 449  
QY 287 RALLQKYADIDIRQDNKTALYWAKEGNATWVRDILQCNPDTEICTKGETPLIKATKM 346  
Db 450 KYLLQNKAKANAKADQDPLHCAARIGHTGMVKLLLENGASPNLATTAGHTPLHTAARE 509  
QY 347 RNIEVVELLDKAKVSAVDKGDTPHLHIAIRGRSRKLAELLILRNPKDGLLYRPNKAGE 406  
Db 510 GHVDTALALEKASQACMTKGGFTPLHVAARYKGVRLAELLEHDA-----HPNAAG 563  
QY 407 ---TPYNIDCSHQ---KSIL-----TOIFGAR----- 427  
Db 564 NGITPLHVAHNNLDIVKLLLRGSGSPHSPANNGYTPHLIAAKQNIQEVARSLLQYGS 623  
QY 428 -----HLSPTETDGMGLYDLVSSA-----LADIL 452  
Db 624 ANAESVQGVTPHLAAQEGHTEMVALLSKQANGLNKNSGLTPLHLVSGEHLVADV 683  
QY 453 SE--PTMOPPICVG-----LYAQWGSK--SPILKKLED-EMKTFAGQIEPLPQSWL- 501  
Db 684 IKHGVTVDAITRMGYTPHVASHYGNKLVKFLLOHQADVNAKTLG--YSPILHQAQOG 741  
QY 502 -IVFLTLLCGGLLFAFTVHN-----LGIAVLSFLALYIPIFIVYFGGRRE 551  
Db 742 HTDIVILLKNG-----ASFNEVSSNGTTPLAIAKRLGYISVTDVLKVT----- 786  
QY 552 GESNNWVLSRLARHIGYLELLKLMMFVNPPELPEQT-TKALPVRFITFDYNRLLSSVG 610  
Db 787 ----DETSVVLVSDKHRMSYPETVDEILDVSEDEGDELVGSKA-----ERRDSRDVG 834  
QY 611 GETSLAEMATLSDACEREGFLATRLFR-----VPKTED-TQKKKWKKTCLPS 660  
Db 835 EEKELDFVPKLDQVVE-----SPAIPRIPCVTPETVIRSEDDQBSKEYDEDSLIPS 888  
QY 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGFLAVLNCKRTWQVLDL 720  
Db 889 -----SPATETSDNI-SPVASPVHTGLVSP-----WVDARG 919  
QY 721 NSQRKRLHNAASKHLKLSGFMKVLKCEVELMARMAKTIDTSFTQNTQRLVWIIDGLDAC 780  
Db 920 GSRGSRHN-----GLRWIP-----PRTCAAPTRITCRV----- 950  
QY 781 EQDKVQLMDTVRLVPSKGPPIAFASDPHIIKAINQNLNSVLRNSNNGHDMYRN-IV 839  
Db 951 ----KPKLNTTPPL-----AEEGLASRIIALGPTGAQFLSPVIV 987  
QY 840 HLPVFL-NSRGLSNARKFLVTSATNGVPCSDTTGTOEDADRVRVSONSLCEMTKLSKTA 898  
Db 988 EIPHFASHRG-----DRELVLVRENSV-----WKEHKSRYGSYLDQILN-GWDE 1035  
QY 899 LNRDRYRRQMQRTITRQMSFDLTKLVV---EDWFSIDISPTQMRLLINVSIVTQRLR 955  
Db 1036 LGSLELEKRVKRCRIIT--TDFPLFYFVMSRLQCD-YDTIGPEG-----GSLRSKLVP 1085

QY 956 ANQISFNWRDLASWINLTQWPYRTSWLILYLEETEGIPDQMTLKIYERISKNIPPTKD 1015  
Db 1086 LVQATEFENAVTKVKLALQ-----AQPVEDLVTKLL----- 1118  
QY 1016 VEPLEIDGDIRNFVEFLSRTPVLVARDVKVFLPCTVNLDPKRLRIIADVRAAREQISI 1075  
Db 1119 -----GNQATF-----SPI-----VIVEPERRKFRPI----- 1141  
QY 1076 GGLAYPPLPLHEGPPRAPSGYSQPPS---VCSSTSPNGFPAGVSPQSHSSYSYSGMTGP 1132  
Db 1142 -GLRIPLPSPWTDNRP-DSGEGDTTSLRLCS-----VIGTDQAO-----WEDITGT 1187  
QY 1133 QHPFYNR-----GSGPAPGVVLLNSL---NVDACEKLL-----QIEGL 1169  
Db 1188 TKLIYANECANFTNVVSARFWLSDCPTABAVHFATLLYKELTAVPYMAKFFVIFAKNDA 1247  
QY 1170 DOSMLPYQCTTIKK-----ANINGRVLAQCNIDELK 1200  
Db 1248 REGRLACYCMTDDKVDKTLLEQHNFVEVARSRDIEVLGMLPFAELSG-----NLVPVK 1301  
QY 1201 KEMNNFQDWHLFRSTVLEMRNAESHVVPEDPRFLSESSGPGAPGEPAPRASHNE----- 1256  
Db 1302 KAAQQRSPHFQSFRENRLAI-----PVKVRDSSREPFGFLSLRKTMYEDTOH 1350  
QY 1257 -LPHTELS-----SOTPYTLNFSFEELNTLGLDEGAPR-----HS 1290  
Db 1351 ILCHLNTMPPCTKGGABDRRTLLPLRLYSILSESRLGFTSDTDRVEMNAVIREHL 1410  
QY 1291 NLSWQSQTR-----RTPSLSSLSNQDSSI-----EISKLTDKVOAEYRD 1329  
Db 1411 GLSWAELARELQFVEDINRIVENPENSLLDQSTALLVWDREGENAKM-ENLYTALRN 1469  
QY 1330 AYREYTAQMSQLEGGPGSTTISGRSPHSTYTYMGQSSSGSIHNSLEQEKGDSEPKPDD 1389  
Db 1470 IDRSEIVNM--LEG-----SGRQS-----RNLKPERRH 1495  
QY 1390 GRKSLMKRGDVIDYSS-----SGVSTNDASPLDPIITEED 1424  
Db 1496 GOREYSLSPQVNGYSSLODELLSPASLOYALPSPLCADQYVNEVTVDAIPL-AAEYHD 1554  
QY 1425 ---EKSDQS--GSKLLPGKKSSESSSLFOT-----DLKLKSGSLRYQKLP-----SD 1466  
Db 1555 TWLEMSDMQVMSAGLTPSLVTAEDSSLECSKAEDSDAIPEWKLEGAHSEDTOGPELGSOD 1614  
QY 1467 EBSGTEESDNT-----PLKDDKDKAKGVKERVKPSPHSABPIRTFFAKAYELSDALL 1522  
Db 1615 LVEDDDTVDSATNGLADLLGOEGORSEKKQREV-SGTQDTEVSLVSGQORVHARIT 1673  
QY 1523 DKK-----DSSDSGVRSSSESPNLSHNEVADDSQLEKANLIELEDSDSHSGKGI--- 1572  
Db 1674 DSPSVQVLDROARQATLWDKQGSTAVHPQEAQSSWQE-----EVTQGHSPFORRITTIQ 1729  
QY 1573 ---PHSLSGLOPFIITARMSICSEDKKSPSECSLIASSPEENWPACQKAYN 1619  
Db 1730 GPPEGALQVEQVIV---STREHVQRPETGTPKAGKEPSLWAPESAFS 1776

## RESULT 12

T42713

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T42713

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, L.

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

the repeat domain

A:Reference number: Z22237; MUID:95340633; PMID:7615634

A:Accession: T42713

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1943 &lt;P&gt;

A:Cross-references: UNIPROT:Q61307; EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.



A: Experimental source: strain C57BL/6J; kidney

C: Geneticks:

A: Gene: Ank3

A: Map position: 10

A: Introns: 855/1

C: Junction:

A: Description: supposed to play an important role in the polarized distribution of many

A: Note: major kidney ankyrin

C: Superfamily: ankyrin; ankyrin repeat homology

C: Keywords: alternative splicing

Query Match 5.9%; Score 523; DB 2; Length 1943;

Best Local Similarity 20.2%; Pred. No. 1.1e-20;

Matches 437; Conservative 270; Mismatches 660; Indels 800; Gaps 90;

```
QY 14 EEENIPALKALEKCKVDNERECQQTPLMTAAEGNLEIVKELIKNGANCNLELDNWT 73
DB 66 KEGHEVVESELLOREANVDAATKGNLTALHTASLAGQAEVVKLVNAGNVAQSQNGFT 125
QY 74 ALISASKEGHVHIVBEELKCGVNLHEDRMGWTALMWCYKGRDVTVELLL----- 124
DB 126 PLYMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQQGHQVVSLLLENDTKGVR 185
QY 125 -----SHGANPSVTGL----- 135
DB 186 LPALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNINVTLLNRAAAYD 245
QY 136 ---QVSVPYIIWAAGRHADIVHLLQNGAKVNCSDKYGTTPLVWAARKG----- 182
DB 246 FTARNDIITPLHVASKRGNANVKKLLDRGAKIDAKTRDGLTPLHCGARSQHEQVEMLLD 305
QY 183 -----HLECVKHLAMGADVDDQEGANSMTALIVAVKGYTQS 219
DB 306 RSAPILSKTKNGLSPLHMAATQDHLNCVQLLQHNVPVDDVNDYLTALHVAACHGVK 365
QY 220 VKEILKRNPNVNLTKGNTALMASKE----- 247
DB 366 AKVLLDKKASPNAKALNGFTPLHIAACKNRIRVWELLKKGASIQAVTESGLTPIHVAAF 425
QY 248 -GHTEIVQDLDAGTYNIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDRGDNKTA 306
DB 426 MGHVNIYVQLMHGASPNNTVRGETAHLMAARSQAEVVRVLDQGAQVEAKAKDQTP 485
QY 307 LYMAVEKGNATWRDILQCNPDTEICTKDGTEPLIKATKMNIEVWELLDKGAKVSADV 366
DB 486 LHSARLGKADIYQQLQOGASPNAAATSGYTPHLHAREGHEDVAAFLDHGASLSITT 545
QY 367 KKGDTPLHIAIRGRSRKLAELLR-----NPKDGRLL----- 398
DB 546 KKGFTPLHVAAYKGLKVASLLOKASPDAAKSGSLTPLHVAAYHYNQKVALLLDQGA 605
QY 399 --YRPNKAGETPYNIDC-SHOKSILTQI--FGAR-----HLSPTETDGMILGY 441
DB 606 SPHAAKNGYTPHLHIAAKNQNDIATSLLEYGADANAVTROGIAVHLLAAQEGHVDVSL 665
QY 442 DLYSALADILSEPTMQB-----PVCGLVLAQM 469
DB 666 LLSRNVNLSNKSGLTPLHIAAQEDRVNVAEVLVNOGAHVDAQTKMGYTPHLVGCH--Y 723
QY 470 GSGK--SFLK--KLEDEMT-----FAQQQ-----IEPLFQFSMLIVFTLLLOGGL 513
DB 724 GNIKIVNELLQHSARVNAKTNGYVHTALHQAQOQGHITHIINVLQNN----- 769
QY 514 GLLFAFTVHPNLGTAVSLFALLYIFPIVIYFGRRREGESNNAWVLSLHARHIGVLE 573
DB 770 ASPNELTVNGTALAI-----ARRLGYS 793
QY 574 LLLKLMFY-----NPPLEPQTTKALPVRFLETDYNRLLSVGGETSIAEMIA 620
DB 794 VVDTLKVVTIEIMTTTTITEKHKNVNPETMNEVLD---MSDDEVKASAPKELSDGEVI- 849
QY 621 TSDACERE-----FGF-LATR--LFRVFKTED--T 646
```

```
DB 850 --SDGECEGDAITGDTDKYLGFDQLKELGDSLPAGYVGFSLGARSASLRSFSSDRSYT 907
QY 647 QGKKKWKTKCCLPSFVIFLPIIGCIITLAIIFRVDP-KHLT-----VNAVLIISI 697
DB 908 LNRSSYAR---DSWHEELVPSKEQHLTTTRFDSLSLHYSWAADTLONVNLVSSPV 963
QY 698 ASVVGLAFVLNCRTWQVLDLNLNQRKRLHNAAS-----KLHLK- 737
DB 964 HSGFLVSMVDARG-----GSMGRSRHGMRIIPPRKCTAPTRITCRLVKRHKLA 1014
QY 738 -----KSEG-----FMKVLKCEVELMARMAMAKTIDSTFQTN 766
DB 1015 NPPPMVEGEGLASRLVEMGPAGQFLGPVIVEIPHFGSMRGKRELIIVLRSENGETWEKH 1074
QY 767 Q-----TRLVIIIDGLDACEODKVLQMLDTRV--LFSKGPFIIFASDPHIILKAINON 819
DB 1075 QFDSKNEDLAELNGMD--BELDSPBELGTVCRIITK-DPPQVFA-----VVSRIKQE 1126
QY 820 LNSVLRSNINGHDMYRNIVHLPVFLNSRGLSNARKFLV--TSATNGDVPCSDTTGIIQ--- 875
DB 1127 SNQIGPEGGI-----LSSITVPLVQASFPEGALTKRIRVGLQAP 1166
QY 876 --EDARRVSONSGEMTKLGSKTALNRDITY--RRQMQRTIITRQMSFDLTKLL---VT 928
DB 1167 VPEETVKKI-----LGNKATFSPITVTEPRRRKFKHPIT--MTPIPVPPSGEGVS 1214
QY 929 EDWFSDISPQTRMLLNIIVSVTGRLLRANQISFNWDRLASWINLQEQW-----YR 979
DB 1215 NGYKGDATP-NLRLUCSITGTSP-----AQWEDITGTTPLTFIKDCVSGFT 1259
QY 980 TS-----WL-----ILYLEETEGIPDQWTLKTI-YERISKNIPTTKOPELLEIDGDIR- 1027
DB 1260 TNVSARFWLADCHQVL---ETVGLASQLYRELIQVPMKFPVFAKTNDP---VSSLRUC 1313
QY 1028 -----NPEVFLSSRTPVLVARDVKVFLPCTVNLDPKUR---ELIADVRA 1068
DB 1314 FCMTDDRVDKTLQEQNEFEVARSK-DIEVLEKPIYVDCYGNLAPLTKGQQLVFNFS 1372
QY 1069 AREQISIGGLAYPLPLHEGPPRAPSGY-----SOPPSVCSSTSFNGFPAGGVSPQPHS 1123
DB 1373 FKEN-----RUPFSIKRDTSQEP--CORLSF----- 1397
QY 1124 SYSGMTGPQHPFYNGSGPAGPFWLNSLVNDAVCEKLRQIQLGDSMLPOYCTTIKK 1183
DB 1398 -----LKEPTKTKGLPQTAV----- 1412
QY 1184 ANINGRVLACNIDELKEMMNFGDWHLPSTVLEMRNAESHVVDPDRFLSSSSSGPA 1243
DB 1413 CNLNLITPLAKKAERKARRQS-----FASLALRKYS-----YLTEPSMSFPQ 1454
QY 1244 PHGBPARR-----ASHNELPHTLSSQTPYTLNFSFEELNTLGLDEGAPRHSNLS-- 1293
DB 1455 --SPCERDIRMAIVADHGLSLWTELARE-----LNFVSDEINQIRVEN---PNSLISQS 1504
QY 1294 -----W---QSTRTPTSSLNSQDSSIBISKLTKVQAEY-----RDAYREYIAQM 1338
DB 1505 FMLLKKVTRDGKNATTDALTSVLTKINRIDIIVLLEGPIDYGNISGTRSFADENNVFH 1564
QY 1339 SOLBGGPG-----STTISGRSSPHSTYMGQSSSGSGSIHS--NL 1375
DB 1565 DPVDGHFSFQVELETPMGLYWTTPNPFQDDHFDSDISSIESPFRTPSRSLDGLVPSQNI 1624
QY 1376 EQEKG-----KDS---EPKPDGKRSFLMKRGDVIDYSSSGSVSTNDASPLDPTTEDE 1425
DB 1625 EHPTGGPPVTAEDTSLSDSKDD-----SVTTVDPADPLD--VDESQ 1665
QY 1426 KSDQSGSKL-----LPGKKSSESS--LFQTDILKXGSLRYQKLPSEDESGTE--- 1473
DB 1666 LKDLQCSECAQCWASVFGIPNDGQAEPLRPQTRKVGMSSEQEKSGKSGPDEEVEDKVK 1725
QY 1474 -----ESDNTPLLKDDKDKAKGKVRVP-----KSPEHSAB-PTRTFIKA 1513
DB 1726 SLFEDIQLEVEABEMTEDQQAMLNRVQRAELAMSSLAGWQNETPTSGSLESQAQ--- 1781
```

QY 1514 KEVLSALLDKDSSDSGVRS-----ESSPNHSLHNEVADDSQLEKANLIEL 1561  
Db 1782 -RLTGGLLDRLODSSDQARDSTSVLTGPGEKIEANGNHTA--EVIPEA---KAKPYFP 1835  
QY 1562 EDDSHSGKRG-----PHLSGLODPIIARMSICSEDKSPSECS--LIASSPENWP 1612  
Db 1836 ESQNDIGKQSIKENLKPHTGCGRTPEV-----SPLTAYQKSLSEETSKLVIEDAPKPCVP 1891  
QY 1613 ACQKAYNLNRPSTVTLNNSAPANQNQFDEMEGIRETSQVILRPPSSPNPTTIQEN 1672  
Db 1892 VGMK--KWTRTTAD-----GKARLNQBEEG-----STRSEPK--QGE 1926  
QY 1673 LKSMTHK 1679  
Db 1927 YKVTXK 1933

RESULT 13  
A55575  
ankyrin 3, long splice form - human  
N;Alternate names: ankyrin G  
C;Species: Homo sapiens (man)  
C;Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 09-Jul-2004  
C;Accession: A55575  
R;Kordeli, E.; Lambert, S.; Bennett, V.  
J. Biol. Chem. 270, 2352-2359, 1995  
A;Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the ax  
A;Reference number: A55575; MUID:95138209; PMID:7836469  
A;Accession: A55575  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-4377 <KOR>  
A;Cross-references: UNIPROT:Q12955; GB:U13616; NID:g608024; PIDN:AAA64834.1; PID:g608025  
C;Genetics:  
A;Gene: GDB:ANK3  
A;Cross-references: GDB:424503; OMIM:600465  
A;Map position: 10q21-10q21  
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology  
C;Keywords: alternative splicing; peripheral membrane protein  
F;73-105/Domain: ankyrin repeat homology <AN01>  
F;106-138/Domain: ankyrin repeat homology <AN02>  
F;139-171/Domain: ankyrin repeat homology <AN03>  
F;172-200/Domain: ankyrin repeat homology <AN04>  
F;201-233/Domain: ankyrin repeat homology <AN05>  
F;234-266/Domain: ankyrin repeat homology <AN06>  
F;267-299/Domain: ankyrin repeat homology <AN07>  
F;300-332/Domain: ankyrin repeat homology <AN08>  
F;333-365/Domain: ankyrin repeat homology <AN09>  
F;366-398/Domain: ankyrin repeat homology <AN10>  
F;399-431/Domain: ankyrin repeat homology <AN11>  
F;432-464/Domain: ankyrin repeat homology <AN12>  
F;465-497/Domain: ankyrin repeat homology <AN13>  
F;498-530/Domain: ankyrin repeat homology <AN14>  
F;531-563/Domain: ankyrin repeat homology <AN15>  
F;564-596/Domain: ankyrin repeat homology <AN16>  
F;597-629/Domain: ankyrin repeat homology <AN17>  
F;630-662/Domain: ankyrin repeat homology <AN18>  
F;663-695/Domain: ankyrin repeat homology <AN19>  
F;696-728/Domain: ankyrin repeat homology <AN20>  
F;729-761/Domain: ankyrin repeat homology <AN21>  
F;762-794/Domain: ankyrin repeat homology <AN22>  
F;795-827/Domain: ankyrin repeat homology <AN23>

Query Match 5.9%; Score 521; DB 2; Length 4377; ~  
Best Local Similarity 20.6%; Pred. No. 5.6e-20;  
Matches 427; Conservative 284; Mismatches 695; Indels 668; Gaps 87;  
QY 14 EENIPALKKCKDKVDVERNECGGTPPLMAEQGNLEIVKELIKNGANCNLEDLDNWT 73  
Db 83 KEHVEVVSSELLQREANVDAATKKGNTHALIASLAGQAEVVKVLVTNGNVAQSQNGFT 142  
QY 74 ALISASKEGHVHVEELLKGVNLEHRDMCGWTALMWACYKGRDVTVELLL----- 124

Db 143 PLYMAAQENHLEVVVKFLLDNGASQSLATEDGFTPLAVALQOQHDOVVSLLENDTKGKVR 202  
QY 125 -----SHGANPSVTGL----- 135  
Db 203 LPALHIAARKDDTKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINVTALLNRAAAVD 262  
QY 136 ---QYVYPIIWAAGRGHADIIVHLLQNGAKVNCSDKYGTTPLVWAARG----- 182  
Db 263 FTARNDDITPLHVASRKGANVMKLLDRGAKIDAKTRDGLTPLHCGARGSEHQVEMVLD 322  
QY 183 -----HLECVKHLAMGADVQDEGANSMTALIVAVKGGVTOS 219  
Db 323 RAAPILSKTYNGLSPLHMAITQGDHLNCVQLLQHNVPDVTNDYLTALHVAACHYKV 382  
QY 220 VKEILKRNPNVNLTKDGNLTALMASKE----- 247  
Db 383 AKVLLDKKANPAKALNGFTPLHIAACKNRIKVMELLKHGASIOAVTESGLTPIHVAAF 442  
QY 248 -GHEIVQDLDAGTYVNIIPDRSGDTVLIGVRGCHVEIVRALLQKVAIDIRQDNKTA 306  
Db 443 MGHVNIIVSQLMHGASPNNTNVRGETALHMAARSGQAEVVRVLVODGAQVEAKADQTP 502  
QY 307 LYWAVEKGNATVVRDILQ--CNPD-----EICT 333  
Db 503 LHSARLGKADIVQQLQOQASPNATTSYTPHLHSAREGHEDVAFLDHGASISITT 562  
QY 334 KQGETPLIKATKMRNIEVVVELLDKGAQSAVDKGGDTPLHIAIRGRSRKLAELLRNPK 393  
Db 563 KKGFTPLHVAAYKGLVANLLLOKASPDAAAGKSLTPLHVAAHYDNOQKALLLL--D 619  
QY 394 DGBLLYRPNKAGETPNIDC--SHQKSLTQI--FGAR-----HLSFTEDGDM 438  
Db 620 QGASPHAAAKNGYTPHLHIAKXNQMDIATLLEYGADANAVTRQGIASVHLAAQEGHVD 679  
QY 439 LGVDLYSSALADITLSEPTQW-----PICVGLY 466  
Db 680 VSLLLGRNANVLSNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVDAQTQNGYTPHVGCH 739  
QY 467 AQWGSCK--SFLIK-KLEDEMTFAQOIEPLQFQSWLIVFLTLCCG-----LGLLFAF 519  
Db 740 --YGNIKIVNLLQHSKAVNAKTNG--YTPLHQAQ-----QGHTHIINVLLQN 785  
QY 520 TWHPN-----LGIIVSLFLALLYIFFYIFGRRREGESWNAWVLSRLARH-I 569  
Db 786 NASPNELTVNGNTALGIARRLGYISVVDTLKIVT-----EETMTTITVTBKHM 834  
QY 570 GYLELLLLKLMFVNPPEL-----PEQTTKALPVRFLFTDYNRLSSV--GGTSLA----- 616  
Db 835 NVPETWNEVLDDSDDEVKANAPE-----MLSDGEYISDVEGEDAMTGDTKYL 884  
QY 617 --EMIAATLSDACEREFGLATRL-----FRVFKTED--TOGKKKWKTCCLPSFVIFLF 666  
Db 885 GPQDLKELGDSLPABGYMGFSLGARSASLSRFSDDRSVTLNRSSVAV--DSMMIEEL 940  
QY 667 IIGCIISGITLAIKFRVDP-KHLT-----VNAVLSIASVUGLAFVNLNCTRWQVLD 717  
Db 941 LVPSKEQHLTTFREEDSDSLRHSWAADTLDNVLNVSPIHSGFLVSFMVDARG----- 994  
QY 718 SLINSORKEHLNAAASKHLKXSEGFMKVL---KCEVELMARMAKTIDSTQNTQRLVVI 773  
Db 995 ---GSVRGSRHNG-----NRILIIPRKCTAP-----TRITCLRV-- 1025  
QY 774 IDGLDACEQDKVQLMDTVRVLFSKGFPIAIFASDPHIIKAINQNLSVLRDSNINGHD 833  
Db 1026 -----KXKLAN-----PPPH-----GERRISSRLVEMGPAGA 1055  
QY 834 YMRN-IVHLPVFLNSRGLSNARKFLVTSATNGDVPCSDTTGTQEDADRRVSQNSLEMTX 892  
Db 1056 FLGPVIVEIPHFSGMRG--KERELIVLRSENGE-----TWKEHQFD---SKNE--DLTE 1102  
QY 893 L--GSKTALNRDVTYRRRQWORTITROMSFDLTLLVTEDFWSDISFQWTRLLNIVSVT 950

Db 1103 LLNGWDELDSPDELGGKRIICRIITKOF-----POYFAVVS--RIKQENQIQPE 1150  
Qy 951 GRLLAAN-----QISFNWDRLASWINLQEQWYRTSMWLYLEETEGIPDQWTLKTIYER 1005  
Db 1151 GGLISSTVPLVQASFPFGALTKIRVGLQ-----AQVPDEIVVKILGNK 1196  
Qy 1006 ISKNIPPTTKDVEPLEIDGDIRNFEVFLSSRTPVLVARDVKVFLPCTVNLNPKLREIAD 1065  
Db 1197 AT-----FSPIVTVPRRKHFKPITMTIPV-----1222  
Qy 1066 VRAAREQISGLLAYPPLPLHEGPPRAPSGYSQP--PSVCSSTSPNGFPAGGVSPQHS 1123  
Db 1223 -----PPSSEGVNSGYKGGDTTNLRLCSIT-----GGTSP-----1254  
Qy 1124 SYISGMTGPHFPYNRGSGPAPGVVLLN-----SLNVD-----AVCEKLKQIEGLDQSM 1173  
Db 1255 AQWEDIQTT-----PLTFIKDCVSFTTNVSARFWLADCHQVLETGLATQL 1301  
Qy 1174 -----LPOYCTTIKKANINGRVLAQCNI-----DELKK--EMMNFQDWHLFRSTVLEM 1220  
Db 1302 YRELICVPYMAKFVVFAMKNDPVESLRCFCWTDKVDKTLQEQENFEE-----VA 1352  
Qy 1221 RNAESHVVPDRPRLSESSSGPAPGEPARPA-----SHNELPHT-----ELSSQTPYT 1269  
Db 1353 RSKDIEVLEGPIYV--DCYGNLAPLTKGGQQLVFNFYFKENRPLFPFKIRDTSQEPCGR 1411  
Qy 1270 LNFSGEELNTLGLDEGAPRHSNLWSQSTRTPSLSSLNQSDSISKLTDKVQAEYRD 1329  
Db 1412 LSLFKERTTKGLQTAVCNINILPAHKKTES-----DQDDIEE-----KTRDQSPASL 1463  
Qy 1330 AYRE-----YIAQMSOLEGGPGSTTISGRSPSHSTY-----YMGQSSGGSIHNSLE 1376  
Db 1464 ALRKRYSLTEPFGMIERTGTAT-----RSLP--TTYSYKPPFFSTRPQSWTTAPITV-----1513  
Qy 1377 QEKGDSEPKDDGKSKFLMRGVDVYSSSGVSTNDASPLDPI-----TEDEKSDQSGS 1432  
Db 1514 -----PQPAKSGF-----TSLSSSSSNTPSASPLKSIWSVSTPSPKISTLGAS 1556  
Qy 1433 KLLPKKSSERASSLFOTDLKLGGLR---YQKLPSEDESGT-----EESDNTPL-----1480  
Db 1557 TTSSVKSISDVASPIRS--LRMTSPKIVTVSQSPNTQVSGTGLARAPVATEPFLKGLA 1615  
Qy 1481 -----LKDKDRKAEGKVER-----VPKSPE-----HSAEPITRFI--KAKELYSDAL 1521  
Db 1616 SNSTFSKRTSPVTTAGSLRLSSITMTTPASPKNINMYSSSLPFPKSIITSAAPLISSPL 1675  
Qy 1522 -----LDKDDSDSGVRSESSPNHSL--HNEVA-----DDSOLEKAN-----LIE-- 1560  
Db 1676 KSVSPVKSRVDVIVSSAKITWASSLSPPVKQMPGHAEVALVNGSISPLKYASSSTLINGC 1735  
Qy 1561 -----LEDDSHSGKGIPIHSLGLODPTIARMSICSEDKKSPSCSLIASSPEENWPAQC 1615  
Db 1736 KATATLQEKISSATNSVSVSAATDVKEKVFSTTTAMPFPLR--SVYSAP-----1786  
Qy 1616 KAYNLNRTPTVTTLNNSAPANRANQNFDEMEGIRETSQVILRPSS-----PNPTTION 1670  
Db 1787 SAFQSLRTPSASALVTSIGSSISATTS-----SVTSSIIITVPVSVVNVLPALPKKL 1839  
Qy 1671 ENLKMTHKRSQRS--YTRLKDDPPELHAASS 1702  
Db 1840 PDSNSFTKSAALLSPIKTLTTTETHPOPHFSRTS 1873

RESULT 14  
T42714  
ankyrin 3, splice form 2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T42714  
R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, I.  
J. Cell Biol. 130, 313-330, 1995  
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene  
the repeat domain.

A:Reference number: Z22237; MUID:95340633; PMID:7615634  
A:Accession: T42714

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1765 <P>

A:Cross-references: UNIPROT:061307; EMBL:L40632; NID:G710548; PID:G710551; PIDN:AA01605

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 5.8%; Score 516; DB 2; Length 1765;

Best Local Similarity 19.8%; Pred. No. 2.4e-20;

Matches 408; Conservative 253; Mismatches 636; Indels 762; Gaps 83;

Qy 14 EEENIPALKLEKCKVDNERNECQGTPLMTAAEQGNLEIVKELIKGANGCNLEDLQNW 73

Db 66 KEGHVEVVSSELLQREANVDAAATKKGNTALHTASLAGQAEVVKVLVTNGANVNAQSQNGFT 125

Qy 74 ALISASKEGVHIVVEELLKCGVNLHEDMGWNTALMACYKGRDVTVELLL-----124

Db 126 PLYMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQQHQVWSLLLENDTKGVR 185

Qy 125 -----SHGANPSVTGI-----135

Db 186 LPALHIAARKDDTAAALLQNDTNADVESKSGTTPHIAAHYGNINVTALLNRAA 245

Qy 136 ---QYSYVPIIWAAGRHADIHLLQNGAKVNSDKYGTTPPLVMAARKG-----182

Db 246 FTARNDTPLHVASKRGKGNANVKKLLDORGAKIDAKTRDGLTFLHCGARSQHEQVVM 305

Qy 183 -----HLCVKKHLLAGADVDBEGANSMTALIVAVKGYTQS 219

Db 306 RSAPILSKTKNGLSPLHMTAQGDHLNCGVQLLQHNVPDVTNDYLTALHVAACHG 365

Qy 220 VKELTKENPNVNLTDKGNLTALMTASKE-----247

Db 366 AKVLLDKKASPNAKALNGFTPLHIACKNIRVWELLKHGASIQAVTESGLTPIHVA 425

Qy 248 -GHTEIVQDLDDAGTYVNIIPDRSGDTVLIGAVRGHVEIVRALLQKYADIDIRQDN 306

Db 426 MGHVNIYSQLMHGCASPNTNVRGETALHMAARSGAEVVRVLYVDQGAQVEAKAKD 485

Qy 307 LYNAVEKGNATWRDIQCNPDTEICTKDGETPLIKATKRNIEVVELLLDKGAKVSAD 366

Db 486 LHISARLGKADIVQQLLQOGASPNAAATTSGYTPLHAAAREGHEDVAAFLPDHGA 545

Qy 367 KKGDTPLHIAIRGSRKSLALLRNPKDGRLLYRPNKAGETPPYNI DCSHQSILTO 426

Db 546 KKGFTPLHVAARKYKLEVASLLQKAS-----PDAAG-----KSGLTPHVA 588

Qy 427 RHLSPTEITDGMGLYDILYSALADILSEPTMQPPICVGLYAQMGSFKLLKLEDEM 486

Db 589 AH-----YDNQKVAL-----LLLDQGASPHAAK-----612

Qy 487 FAGQIBPLFQFQSWLIVFLTLCCGGLLPFAFTVHPNLGIANVLSLALLIYFIV 546

Db 613 -----NGYTPLHIAAKNQMDIATSL-----LEY 636

Qy 547 GG-----RREGESNNAWVLSLRLARHIGYLELLLLKLMFVNPPELPBQTKALPV 601

Db 637 GADANAVTQG-----IASVHLAAQEGHVDWMVLSLSRN-----A 671

Qy 602 DYNRLSSVGGETSIAEMIATLSDACEREFGLFRLFRVFKTE-----DTQGGKKW 652

Db 672 NVN-LSNKSGLTLP-----HLAAQEDRVNVAEVLVNOGAHVDAQTQMGY 714

Qy 653 KKTCCLPSPFVIFPIFCIISGITLLAIFRVDPKHLTVNAVLIASVGLAFVLCRTW 712

Db 715 TP-----LHVCHYGNIKI-----VNFLLQHSAKV-----NAKT- 743



```

QY 199 -----NSMTALIVAKGTYQSVKEILKR 226
Db 338 ITLOHITPLHVAACHGHRVAKVLLDKCAKPNRSLNGFTPLHIAACKKHVRWELLKLT 397
QY 227 NPNVNLTKDGNATMIASKEGHEIIVODLLDAGTYVNIPIRSGDTVLIGAVRGHVEIV 286
Db 398 GASIDAVTESGLTPLHVASFMGLPIVKNLLQRGASPNVSVKVTETPLHMAARAGHTEVA 457
QY 287 RALLQKYADIDIRGQDNATLYAVAKGNATWNRDILQCNPDTEICTKDGTEPLIKATKM 346
Db 458 KYLLQNKAKAKAKDDOTPLHCAARIGHTWKNLLENGASPNLATTAGHTPLHTAARE 517
QY 347 RNIEVELLLDKGAKVSADVKGTPLHIAIRGRSRKLAELLRNPKDGRLLYRPKNAGE 406
Db 518 GHVDITALALLEKEASQACWTKKGTPLHVAAYKGVRLAELLEHDA-----HPNAAGK 571
QY 407 ---TPYNDCHQ-----KSIL-----TQIFGAR----- 427
Db 572 NGLTPLHVAVHNHNDIIVKLLPRGSGSPHSPAWNNGYTPHIAAKQNIQEVARSLLQYGG 631
QY 428 -----HLSPTEDGMLGYDLYSSALADILSEPTMQP-----PIC--- 462
Db 632 ANAESVQGVTPHLHAAQEGHTEMVALLSKQANGNLGNKSLTPLHLVSEQGHVFPADV 691
QY 463 -----VGLYAQMGSGK--SFLAKKLED-EMKTPAGQOIEPLFQFSL- 501
Db 692 IKHGVTVDAATRMGYTPHVAASHYGNIKLVKFLHQHQAQVNAKTKLG--YSPLHQAQOG 749
QY 502 -IVFTLLCGGLGLLFAFTVHPN-----LGIAVSLFLALLYFFIVYFGGRRE 551
Db 750 HTDVTLLKNG-----ASPNEVSSNGTTPLAIAKRLGYISVTDVLKVV- 794
QY 552 GESNNAWVSTRLARHIGYLELLKLMFVNPPPEQTITKALPV--RPLFTDYNRLSS 608
Db 795 ----DETSVLVSDKHMSYPETVDEILDVS-----EDEGTAHISIMGDELVGSKAERDS 846
QY 609 --VGGETSLAEMIAATLSDACREFGFLATRFR-----VEKTED-TGCKKWKKTC 656
Db 847 RDVGEERKELDFVKLDQVVE-----SPAIPRIPCVTPTVIRSEDOEASKEYEDS 900
QY 657 CLPSFVIFLTIIGIITLAILFRVDPKHLTVNAVLIISIASVGLAFVLCNCTWMOVL 716
Db 901 LIPS-----SPATETSDNI-SPVASPVHTGFLVSF-----MV 931
QY 717 DSLNSOKRHLHNAASKLHLKSGFMKVLKCEVELMARMAKTIDSPQNTQTRLVIIIDG 776
Db 932 DARGGSMRGRHN-----GLRVIP-----PRCAAPTRITCRLV----- 966
QY 777 LDACEQDKVLQMLDTRVFLSKGPPIAFASDPHIIKAIINQNLNSVLNDSNINGHDYMR 836
Db 967 -----KPQKANTPPPL-----AEEGLASRIIALGPTGAQFLS 999
QY 837 N-IVHLPVFL-NSRGLSNARKFVTSATNGVPCSDTTGIEDADRRVSONSLGEMTKLG 894
Db 1000 PVIIEIPHAFASHGRG---DRELVLRSNGSV-----WKEHKSRYGESYLDQILN-G 1047
QY 895 SKTALNRDTRRRQMORTIIFROMSFDLTKLVT-----EDWFSDISPQTMRLNINVSUTG 951
Db 1048 MDELSGLEEKRVCRKRIIT--TDFPLYFVINSRLOOD-YDTIGPEG-----GSLRS 1097
QY 952 RLLRANQISFNWDRLASINILTEQWPYRTSMILYLBETEGIPQOMTLKTIYERISKNI 1011
Db 1098 KLVPVQATFFENAVTNKVKLALQ-----AQVPDELVTKLL----- 1134
QY 1012 TTQDVEPILLEIDIRNFEVFLSSRTPVLVARDVKVFLPCTVNLDPKLRLEIADVRAARE 1071
Db 1135 -----GNOATF-----SPI-----VTVEPRRKRPHRI----- 1157
QY 1072 QISIGGLAYPLPLHEGPAPRAPSQPPS---VCSSTSFGNGPPAGGVSVOPHSSYSYG 1128
Db 1158 ----GLRIPLPSPWTONPR-DSGEGDTTSLRLCS-----VIGGTDQAQ-----WED 1199

```

```

QY 1129 MTGPQHFPYNR-----GSGPAPGPVVLNLSL---NVDVAVCEKJK-----Q 1165
Db 1200 ITGTTKLIYANECANFTTNSARFWSLSDCPTAEAVHFATLLYKELTAVPYMAKFVIFAK 1259
QY 1166 IEGLDQSMPLPOYCTTIK-----ANINGRVLAQCN 1196
Db 1260 MNDAREGLRCYCMTDDKDKVKTLEQHENFVEVARSRDIEVLEGMPLFAELSG-----NL 1313
QY 1197 DELKEMNMFGDWHLFRSTVLEMRNAESHVVVPEDPRFLSSSSGCPAPHGEPARRASHNE 1256
Db 1314 VPVKAQAQSRFHFQSFRENRLAI-----PKVRSRSREPFGGFLFLRXTMYE 1362
QY 1257 ----LPHTELS-----SQPYTLNFSFEELNTLGLDEGAPR----- 1288
Db 1363 DTQHILCHLNTMPPCTKSGAEDRRRTLTPLTLRYSLTSSRLGFTSDTORVEMRMAVI 1422
QY 1289 --HNLWSQSTR-----RTPSLSSINSQDSSI-----EISLTKDKVOA 1325
Db 1423 REHLGLSMAELARELOFVEDINIRVENPNNSLLDQSTALLTWVDREGENAKM-ENLYT 1481
QY 1326 EYRDAYREYIAQMSQLEGPGSTTISGRSSPHSTVYMQSSSSGSIHNSLNQEKGDSEP 1385
Db 1482 ALNRIDRSEIVNMLE-----VSGRQS-----RNKLP 1507
QY 1386 KPDPGRKSLMKRGDVIDYSS-----SGVSTNDASPLDPIERDEKSDQSGSKLLPGKK 1439
Db 1508 ERRHGDREYSLSPSQVNGYSSLODELLSPASLQYALP-SPICADQYNEVTVIDAIPLAA 1566
QY 1440 SSERSSLFQDTLKLKSGSLRYQKLPSEDE---SGTESDNTPLLKDDKDKRAEGKVERV 1496
Db 1567 TEHDTMLEMSDMQVMSAGLTPSLVTAEDSSLECSKAEDSDAIP-----EWKLEGAHSED 1620
QY 1497 PKSPHSAEPI---RTFIKAYEVLSDALLDKK-----DSSDSGVRSSSES 1537
Db 1621 TQGPDELGSQDLVEDDVTVDSDATNGLADLLGQORVHARITDPSVSRQVLDLRSOARTLDWK 1680
QY 1538 SPNHSLENEVADDSOLEKANLIELEDDSHSGKGI-----PHSLSGLODPIIARMISIC 1590
Db 1681 QGSTAVHPQEAQTSQWQE---EVTQGHPSFORRITTIQGPPEGALQEYQVLV---STR 1733
QY 1591 SEDKKSPPSECSLIASSPEENWPACQKAYN 1619
Db 1734 EHVGQPPETGSPKAGKEPSLWAPESAFS 1762

```

Search completed: February 5, 2005, 19:28:18  
Job time : 58 secs

**This Page Blank (uspto)**





QY 241 LMTASKEGHTETIYQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKVDIDIRG 300  
DB 246 LMTASKEGHTETIYQDLDAGTYVNI PDRSGDTVLIGAVRGHVEIVRALLQKVDIDIRG 305  
QY 301 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKDGSTPLIKATKMRNIEVVELLDKGA 360  
DB 306 QDNKTALYWAVERKGNATWVRDILQCNPDTEICTKDGSTPLIKATKMRNIEVVELLDKGA 365  
QY 361 KVSADVKKGTPTPHIAIRGRSKLAELLNPNKGRLLYRPNKAGETPPYINDCSHOKSIL 420  
DB 366 KVSADVKKGTPTPHIAIRGRSKLAELLNPNKGRLLYRPNKAGETPPYINDCSHOKSIL 425  
QY 421 TQIFGARHLSPTTGDGMLGYDLYSSALADILSEPTMOPPICVGLYAQWGSFGKLLKLL 480  
DB 426 TQIFGARHLSPTTGDGMLGYDLYSSALADILSEPTMOPPICVGLYAQWGSFGKLLKLL 485  
QY 481 EDEMKTTFAGQOIEPLFQFQSWLIVFLTLILCGGLGLLFAFTVHNLGIAVSLFLALXYIF 540  
DB 486 EDEMKTTFAGQOIEPLFQFQSWLIVFLTLILCGGLGLLFAFTVHNLGIAVSLFLALXYIF 545  
QY 541 FIVYFGRRREGESWNAWYLSRLARHIGYLELLKLMFVNPPPELPEQTTKALPVRFLF 600  
DB 546 FIVYFGRRREGESWNAWYLSRLARHIGYLELLKLMFVNPPPELPEQTTKALPVRFLF 605  
QY 601 TDYNRLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFKTEDTQKKKWKKTCLPS 660  
DB 606 TDYNRLSSVGGETSLEAMIAATLSDACEREFGLATRLFRVFKTEDTQKKKWKKTCLPS 665  
QY 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGFLAVLNCKRTWQVLDLSLL 720  
DB 666 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGFLAVLNCKRTWQVLDLSLL 725  
QY 721 NSQKRLHNAASKHLKLGSEFGKVLKCEVELMARMAKTTIDSTQNTQRLVVIIDGLDAC 780  
DB 726 NSQKRLHNAASKHLKLGSEFGKVLKCEVELMARMAKTTIDSTQNTQRLVVIIDGLDAC 785  
QY 781 EQDKVLQMLDVTVLVSKGPIIAIFASDPHIIKATNONLNSVLRDSNNGHDMYRNIVH 840  
DB 786 EQDKVLQMLDVTVLVSKGPIIAIFASDPHIIKATNONLNSVLRDSNNGHDMYRNIVH 845  
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCDTTQIEDADRRVQNSLGEMLKLSGTALN 900  
DB 846 LPVFLNSRGLSNARKFLVTSATNGDVPSCDTTQIEDADRRVQNSLGEMLKLSGTALN 905  
QY 901 RRDYTRRRQMRITRQMSFDLTKLIVTEWDFSDISPTQWRRLLNVSVTGRLLRANQIS 960  
DB 906 RRDYTRRRQMRITRQMSFDLTKLIVTEWDFSDISPTQWRRLLNVSVTGRLLRANQIS 965  
QY 961 FNDRLASWNLTEQWPYRTSWILVLEETEGIPDQWTLKTIYERISKNIPTTKDVEPLL 1020  
DB 966 FNDRLASWNLTEQWPYRTSWILVLEETEGIPDQWTLKTIYERISKNIPTTKDVEPLL 1025  
QY 1021 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGGLAY 1080  
DB 1026 EIDGDIRNFVFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGGLAY 1085  
QY 1081 PPLPLHEGPPRAPSGVSVQSPVCSSTSFNGPPAGGVSVQPHSSYSGMTGPOHPYFN 1138  
DB 1086 PPLPLHEGPPRAPSGVSVQSPVCSSTSFNGPPAGGVSVQPHSSYSGMTGPOHPYFN 1145  
QY 1139 -----RGSGP 1143  
DB 1146 FPAPLYTPRYPGGQHLISRSVKTSLPRDQNGLEVIKEDAAEGLSPDTSRGSGP 1205  
QY 1144 APGFVVLLNSLVNDVACEKLKQIEGLDQSLMPOYCTTIKKANINGRVLQACNIDELKEM 1203  
DB 1206 APGFVVLLNSLVNDVACEKLKQIEGLDQSLMPOYCTTIKKANINGRVLQACNIDELKEM 1265  
QY 1204 NMFGDWHLPRSTVLEMRNASHVWPDPRFLSESSGAPGPHGEPARRASHNELPHTELS 1263  
DB 1266 NMFGDWHLPRSTVLEMRNASHVWPDPRFLSESSGAPGPHGEPARRASHNELPHTELS 1325  
QY 1264 SQPTPTLNFPEELNTLGLDEGAPRHSNLSWQSQTRRTTPELSLSLNSQDSSIEISKLTDKV 1323

DB 1326 SQPTPTLNFPEELNTLGLDEGAPRHSNLSWQSQTRRTTPELSLSLNSQDSSIEISKLTDKV 1385  
QY 1324 QAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGSIHNSLBOEKKGDS 1383  
DB 1386 QAEYRDAYREYIAQMSQLEGGPGSTTISGRSSPHSTYYMGQSSSGSIHNSLBOEKKGDS 1445  
QY 1384 EPKPDGRKSFILMKRGDVIDYSSGVSTNDASPLDPIETEDEKSDQSGKLLPGKKSSER 1443  
DB 1446 EPKPDGRKSFILMKRGDVIDYSSGVSTNDASPLDPIETEDEKSDQSGKLLPGKKSSER 1505  
QY 1444 SSLFQDLDLKGSLRYQKLPSEDESGTTEESDNTPLLDKDKRKAEGKVERVPKSPHS 1503  
DB 1506 SSLFQDLDLKGSLRYQKLPSEDESGTTEESDNTPLLDKDKRKAEGKVERVPKSPHS 1565  
QY 1504 ASPIRTFIKAEYLSKDALDKKSDSGSVRSSESSPNHSLHNEVADDSOLEKANLIELED 1563  
DB 1566 ASPIRTFIKAEYLSKDALDKKSDSGSVRSSESSPNHSLHNEVADDSOLEKANLIELED 1625  
QY 1564 DSHSGKRGIPHSLSGLQDDPIIARMSICSEDKKSPSECSLIASSPENWPAQKAYNLNRT 1623  
DB 1626 DSHSGKRGIPHSLSGLQDDPIIARMSICSEDKKSPSECSLIASSPENWPAQKAYNLNRT 1685  
QY 1624 PSTVTLLNNSAPANRANQNFDEMEGIRETSQVILRPSPPNPTTIQENLKSMTKRQSQR 1683  
DB 1686 PSTVTLLNNSAPANRANQNFDEMEGIRETSQVILRPSPPNPTTIQENLKSMTKRQSQR 1745  
QY 1684 SSYTRLKDPPELHAAASSESTGFGBERESIL 1715  
DB 1746 SSYTRLKDPPELHAAASSESTGFGBERESIL 1777

RESULT 2  
Q9EQG6 PRELIMINARY; PRT; 1762 AA.  
AC Q9EQG6;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE KIDINS220.  
GN Name=Kidins220;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20568256; PubMed=10998417; DOI=10.1074/jbc.M005261200;  
RA Iglesias T., Cabrera-Poch N., Mitchell M.P., Naven T.J., Rozengurt E.,  
RA Schiavo G.;  
RT Identification and cloning of Kidins220, a novel neuronal substrate  
of protein kinase D.;  
RL J. Biol. Chem. 275:40048-40056 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Mitchell M., Schiavo G.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF23045; AAG35185.2; -.  
DR HSSP; P09959; 1SW6.  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; Ank; 11.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 11.  
DR PROSITE; PS50088; ANK\_REPEAT; 10.  
DR PROSITE; PS50297; ANK\_REPEAT; 1.  
KW ANK repeat.  
SQ SEQUENCE 1762 AA; 195714 MW; 0CB2689A571F8AE4 CRC64;

Query Match 94.1%; Score 8362.5; DB 2; Length 1762;  
Best Local Similarity 91.6%; Pred. No. 0;  
Matches 1616; Conservative 47; Mismatches 50; Indels 51; Gaps 4;  
QY 1 MSVLISQSVINVEENIPALKLEKCKDVDERNECGQTPLMIAAEQGNLEIVKELKN 60

Db	1	MSVLISQSVINYVEENIPALKVALECKQVDERNECCQTEPLMAAEGQNVIEIVKELLN	60
Qy	61	GANCNLEDLNDWTALISASKEGHVHIVBEELKCGVNLHEDMGWGTALMAWACYKGRDVV	120
Db	61	GANCNLEDLNDWTALISASKEGHVHIVBEELKCGVNLHEDMGWGTALMAWACYKGRDVV	120
Qy	121	ELLLSHGANSPTSGLQSVYPIIWAAGRGHADIVHLLQLONGAKVNCSDKYGTTPPLVMAAR	180
Db	121	ELLLSHGANSPTSGLQSVYPIIWAAGRGHADIVHLLQLONGAKVNCSDKYGTTPPLVMAAR	179
Qy	181	KGHLECVKHLAMGADYDOEGANSHTALIIVAVKGGYTQSVEIILKRPNNVNLTDKGNTA	240
Db	180	KGHLECVKHLAMGADYDOEGANSHTALIIVAVKGGYTQSVEIILKRPNNVNLTDKGNTA	239
Qy	241	LMIASKEGHTIEIVODLLDAGTYVNIIPDRSGDPTVLIGAVRGHVEIVRALQKYADIDIRG	300
Db	240	LMIASKEGHTIEIVODLLDAGTYVNIIPDRSGDPTVLIGAVRGHVEIVRALQKYADIDIRG	299
Qy	301	QDNKTALYAWVEKGNATMVRDILQCNPDTEICTKDGETPLIKATMARNIEVVELLDKGA	360
Db	300	QDNKTALYAWVEKGNATMVRDILQCNPDTEICTKDGETPLIKATMARNIEVVELLDKGA	359
Qy	361	KVSAVDKKGDTPLHAIIRGRSRKLAELLLRNPKGDRLLYRNKAGETPTYNIDCSHQSKIL	420
Db	360	KVSAVDKKGDTPLHAIIRGRSRRLAEILLRNPKGDRLLYRNKAGETPTYNIDCSHQSKIL	419
Qy	421	TOIFGARHLSPTETDGMGLGYDLYSSALADTILSEPTMQPPICVGLYIAQWGSKGKFLKKL	480
Db	420	TOIFGARHLSPTETDGMGLGYDLYSSALADTILSEPTMQPPICVGLYIAQWGSKGKFLKKL	479
Qy	481	EDEMKTFAGQOIETPLFQFWSMLIVELTLLCGGLGLFAFVHPNLGIAVSLSFLLALYIF	540
Db	480	EDEMKTFAGQOETPLFQFWSMLIVELTLLCGGLGLFAFVTDNLTAISLSFLALYIP	539
Qy	541	FIVIFYGCRREGESWNWAWLSTRLAHIGYLELLKLMPVNPPELPEQTTKALPVRFPL	600
Db	540	FIVIFYGCRREGESWNWAWLSTRLAHIGYLELLFKLMPVNPPELPEQTTKALPVRFPL	599
Qy	601	TDYNRLSSVGGETSLAEMIAITLSDACRREFGFLATRLFRVPKTEDTGCKKKWKTCCLPS	660
Db	600	TDYNRLSSVGGETSLAEMIAITLSDACRREFGFLATRLFRVPRTSESGCKKKWKTCCLPS	659
Qy	661	FVIFLFIIGCIITLAIIPRVDPKHLTVNAVLSIASVVGGLAFVNLNCRTWQVLDSLL	720
Db	660	FVIFLFIIGCIITLAIIPRVDPKHLTVNAVLSIASVVGGLAFVNLNCRTWQVLDSLL	719
Qy	721	NSQKRRLHNAASKLHKIKSGGFMKVLCEVELMARMAKTIDSFQONQTRLVVIIDGLDAC	780
Db	720	NSQKRRLHNAASKLHKIKSGGFMKVLCEVELMARMAKTIDSFQONQTRLVVIIDGLDAC	779
Qy	781	EQDKVLOMLDTRVLFPSKGPIAIFASDPHIIKAIINONLSVLRDSNINCHDYMRNVH	840
Db	780	EQDKVLOMLDTRVLFPSKGPIAIFASDPHIIKAIINONLSVLRDSNINCHDYMRNVH	839
Qy	841	LPVFLNSRGLSNARKFLVTSATNGDVPSCSDTTGQEDADRVSQNSLGEKMTKLGSKTALN	900
Db	840	LPVFLNSRGLSNARKFLVTSATNGDITCSDDTTGQEDTDRVSQNSLGEKMTKLGSKTALN	899
Qy	901	RRDYYRRRQMTITRQMSFDLTKLAVTEBWFSDISPOTWRRLLNIVSVTGRLLRANQIS	960
Db	900	RRDYYRRRQMTITRQMSFDLTKLAVTEBWFSDISPOTWRRLLNIVSVTGRLLRANQIT	959
Qy	961	FNNDRLASWNLTEQWPYRTSWLLILYEEETGIPDQMTLKTIVIERISKNTPTTKQVPEPL	1020
Db	960	FNNDRLASWNLTEQWPYRTSWLLILYEEETGIPDQMTLKTIVIERISKNTPTTKQVPEPL	1019
Qy	1021	EIDGDIRNFVEFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQISIGSLAY	1080
Db	1020	EIDGDIRNFVEFLSSRTPVLVARDVKVFLPCTVNLDPKLEIITADVRAAREQINIGSLAY	1079
Qy	1081	PPPLPHEGPPRAPSGYQOPPSVCSSTSFNGPFAAGVVSPPPHSSYSSGMTQPHQPFYNR	1139

Db	1080	PPLPLHGGPPRPSPGYSQPASVCSASFNGFPFGVGVVSPQPHSYSGLSGLSGPQHFPYRNP	1133
Qy	1140	-----GSG-----PAGPVL	1151
Db	1140	FPAPLYTPRYPGCSOHLISRSVKTSLPDQVNGLPDGSFKNQROAAVPATGSLLL	1199
Qy	1152	NSLNVDAVCEKLQIEGLDQSLPQYCTTIKCANINGRVLAQCNIDELKEMMNFQDWH	1211
Db	1200	SSMTVDVCEKLQIEGLDQSMMPQYCTTIKCANINGRVLSQCNIDELKEMAMNFQDWH	1259
Qy	1212	LPFRSTVLEMRNAESHVVPEDPRFLSESSSPAPHGEPAPRASHNELPHTLSQTPTTLN	1271
Db	1260	LFRSKVLEMRSESVQVPEDPRFLNENSAVPVPHGESARSSHTPLTLTSLSSQTPTTLN	1319
Qy	1272	FSFELNTLGLDEGAPRHSNLWSQSQRRTPSPSLNSQDSSSTEISKLTDKVOAEYRDAY	1331
Db	1320	FSFELNTLGLDEGAPRHSNLWSQSQRRTPSPSLNSQDSSSTEISKLTDKVOAEYRDAY	1379
Qy	1332	REYIAQMSQLEGGPGSTTISGRSPHSTYMGQSSSGSITHNLEQEKGDSPKPPDDGR	1391
Db	1380	REYIAQMSQLEGGTGSSTISGRSPHSTYMGQSSSGSITHLEQERKGEGLKQEDGR	1439
Qy	1392	KSPLMKGGDIVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKKSSERSLFTQDL	1451
Db	1440	KSFPMKGGDIVIDYSSSGVSTNEASPLDPIITEDEKSDQSGSKLLPGKKSSERSLFTQDL	1499
Qy	1452	KLKSGGLRYOKLPSEDESGTGESDNTPLKDDKDKAECKVRRVPKSPHSAEPIRTFI	1511
Db	1500	KLKGGGLRYOKLPSEDESGTGESDNTPLKDDKDKAEKGAERVKSPHSAEPIRTFI	1559
Qy	1512	KAKYLSDALLDKKDSSDGVRSSESPNHSNLEHVADDQLEKANLIELEDDSHSGKRG	1571
Db	1560	KAKYLSDALLDKKDSSDGVRSNESSPNHSNLEHVADDQLEKANLIELEDDSHSGKRG	1619
Qy	1572	IPHSLSGLQDPIIARMGICSEDKKSPSECSLIASSPEENVPACQKAYNLNRTPTSTVLNN	1631
Db	1620	MPHLSLSGLQDPIIARMGICSEDKKSPSECSLIASSPEESWPACQKAYNLNRTPTSTVLNN	1679
Qy	1632	NSAPANANQNFDEMGIRTSQVILRPSSSPNPTTIQENLKSMTHKRSQSSSYTRLSK	1691
Db	1680	NTATPNRANQNFDEMGIRTSQVILRPSPNPTAVQENLKSMAHKRSQSSSYTRLSK	1739
Qy	1692	DPPELHAAASSESTGFEERESIL 1715	
Db	1740	DASELH-AASSESTGFEERESIL 1762	
RESULT 3			
Q9ERD4			
ID	Q9ERD4	PRELIMINARY;	PRT; 1715 AA.
AC	Q9ERD4;		
DT	01-MAR-2001	(TrEMBLrel. 16, Created)	
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Ankyrin repeat-rich membrane-spanning protein.		
GN	Name=ARMS;		
OS	Eukaryota norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley;		
EX	MEDLINE=20585245; PubMed=11150334;		
RA	Kong H., Boulter J., Weber J.L., Lai C., Chao M.V.;		
RT	"An evolutionarily conserved transmembrane protein that is a novel		
RT	downstream target of neurotrophin and ephrin receptors.";		
RL	J. Neurosci. 21:176-185 (2001).		
DR	EMBL; AF313464; AAC34167.1; -.		
DR	HSPF; P09959; ISW6.		
DR	InterPro; IPR002110; ANK.		
DR	Pfam; PF00023; Ank; 11.		
DR	PRINTS; PR01415; ANKYRIN.		



"Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous  
 cDNAs identified by screening of terminal sequences of cDNA clones  
 randomly sampled from size-fractionated libraries.";

RT DNA Res. 10:35-48(2003).  
 DR EMBL; AK122478; BAC65760.1; -.  
 DR HSP; P09959; I5W6.  
 DR MGD; MGI:1924730; C330002I19Rik.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR InterPro; IPR002110; ANK.  
 DR Pfam; PF00023; Ank; 10.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 9.  
 DR PROSITE; PS50088; ANK\_REPEAT; 8.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 KW ANK repeat.  
 FT NON TER.  
 SQ SEQUENCE 1 188253 MW; 599DE2A3419D4C9D CRC64;

Query Match 87.5%; Score 7771; DB 2; Length 1693;  
 Best Local Similarity 87.9%; Pred. NO. 0;  
 Matches 1506; Conservative 54; Mismatches 53; Indels 100; Gaps 6;

QY 83 HVHVEELLKGVNLEHDMGWTALMWACYKGRDVVELLSHGANSVTGLQYSYPI 142  
 DB 1 HIHVEELLKGVNLEHDMGWTALMWACYKGRDVVELLSHGANSVTGL-QYSYPI 59  
 QY 143 IWAAGRGHADIVHLLQNGAKVNSDKYGTTPLVWAARKHLECVKHLANGADVDBGA 202  
 DB 60 IWAAGRGHADIVHLLQNGAKVNSDKYGTTPLVWAARKHLECVKHLANGADVDBGA 119  
 QY 203 NSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNALMIASKEGTEIIVQDLLDAGTY 262  
 DB 120 NSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNALMIASKEGTEIIVQDLLDAGTY 179  
 QY 263 VNIPDRSDTVLIGAVRGHVEIVRALLQKYADIDIRQDNKNTALYNAVEKGNATVRDI 322  
 DB 180 VNIPDRSDTVLIGAVRGHVEIVRALLQKYADIDIRQDNKNTALYNAVEKGNATVRDI 239  
 QY 323 LQCPNDETEICTKOGTEPLIKATKRNIEVVELLDKGAQVAVDKKGTPLHIAIRGRSR 382  
 DB 240 LQCPNDETEICTKOGTEPLIKATKRNIEVVELLDKGAQVAVDKKGTPLHIAIRGRSR 299  
 QY 383 KLABELLNPKDGRLLYRPNKAGETPNYDCSHQKSLITQIFGARHLSPTEIDGMLGYD 442  
 DB 300 RLABELLNPKDGRLLYRPNKAGETPNYDCSHQKSLITQIFGARHLSPTEIDGMLGYD 359  
 QY 443 LYSALADILSEPTWQPPICVGLYAQWGSFKLLKLEDEMKTFAQQIIEPLFQFSLWI 502  
 DB 360 LYSALADILSEPTWQPPICVGLYAQWGSFKLLKLEDEMKTFAQQIIEPLFQFSLWI 419  
 QY 503 VFLTLLCGGLGGLFAFTVHNLGTAVSLFLALLYIPFIVYFGRRGEGSWNAWVLS 562  
 DB 420 VFLTLLCGGLGGLFAFTVHNLGTAVSLFLALLYIPFIVYFGRRGEGSWNAWVLS 479  
 QY 563 TRLARHIGYELLKLMFNPPELPEQTTKALPVRFLETDYRLSSVGGETSLEMIATL 622  
 DB 480 TRLARHIGYELLKLMFNPPELPEQTTKALPVRFLETDYRLSSVGGETSLEMIATL 536  
 QY 623 SDACERBFGFLATRLFRVFKTEDTQKKWKTKTCLPSFVIFLPIIGCIIGITLLAIFR 682  
 DB 537 -----FEYLRLLKILRY-----KRNKKTCLPSFVIFLPIIGCIIGITLLAIFR 581  
 QY 683 VDPKHLTVNAVLIISTASVGLAFVNLCKTWQVLDLSLNSQRKHLNAAKHLKSKGFP 742  
 DB 582 VDPKHLTVNAVLIISTASVGLAFVNLCKTWQVLDLSLNSQRKHLNAAKHLKSKGFP 641  
 QY 743 MKVLKCEVELMARMAKTIDSTFQNTRLVLIIDGLDACEQDKVLQMLDTRVLFSGKGPFI 802  
 DB 642 MKVLKCEVELMARMAKTIDSTFQNTRLVLIIDGLDACEQDKVLQMLDTRVLFSGKGPFI 701  
 QY 803 AIFASDPHIITKAINQNLSVLRDSNNGHDMYRNIVHLPVFLNSRGLSNARKFLVTSAT 862

Db 702 AIFASDPHIITKAINQNLSVLRDSNNGHDMYRNIVHLPVFLNSRGLSNARKFLVTSAT 761  
 QY 863 NGDVPCSDTTTGIQEDADRRVVSQNSIGEMTKLGSKTALNRDITYRRRQMQRTITRQMSFDL 922  
 Db 762 NGDISCEATGQEDADRRVVSQNSIGEMTKLGSKTALNRDITYRRRQMQRTITRQMSFDL 821  
 QY 923 TKLLVTEDEWFSDISPQWTRRLNIVSVTGRLLRANQISFNWDRLASINLWTEQWPIYTSW 982  
 Db 822 TKLLVTEDEWFSDISPQWTRRLNIVSVTGRLLRANQITFNWDRLASINLWTEQWPIYTSW 881  
 QY 983 LILYLEETEGIPDOMTKTIYERISKNIPPTKQVEPLEIDGDRNFEVLSRTPVLVA 1042  
 Db 882 LILYLEETEGIPDOMTKTIYERISKNIPPTKQVEPLEIDGDRNFEVLSRTPVLVA 941  
 QY 1043 RDVKVFLPCTVNLDPKLEIIVADVAAREQISIGLAYPPLPLHEGPPRAPSGYSPQPSV 1102  
 Db 942 RDVKVFLPCTVNLDPKLEIIVADVAAREQINIGLAYPPLPLHEAPRPSGYSQPASV 1001  
 QY 1103 CSST-SFNGPPAGGVSPQPHSSYISGWTGQHPFYNR----- 1139  
 Db 1002 CSSSASFGPPGWSVPQPHSSYISGLSGQHPFYNRFPFAPVLYTPRYYPGSGSHLIS 1061  
 QY 1140 -----CSG 1142  
 Db 1142 RSVKTSPLRQNNGLVIEKEDAAEGLPSPFASREKSWTRKQMLCDSGFNKORQASV 1121  
 QY 1143 PAPGVVLLNSLVDAVCEKLRQIEGLDQSMPLPOYCTTIKANINGRVLACNIDELKKE 1202  
 Db 1122 PATGTSLLSSMTVDVCEKLRQIEGLDQSMPLPOYCTTIKANINGRVLACNIDELKKE 1181  
 QY 1203 MNMFGDWHLFRSTVLEMRNABSHVPEDPFRFLSESSGAPHPGEPARRASHNELPHTL 1262  
 Db 1182 MAMNFGDWHLFRSVLEMRNABSHVPEDPFRFLSESSGAPHPGEPARRASHNELPHTL 1241  
 QY 1263 SSQTPYTLNPSFELNLTGLDEGAPRHSNLSWQSTRTPTSLSSINSDSIEISKLTDK 1322  
 Db 1342 SSQTPYTLNPSFELNLTGLDEGAPRHSNLSWQSTRTPTSLSSINSDSIEISKLTDK 1301  
 QY 1323 VQAEYRDAYREYIAQMSQLEGPGSTTISGRSSPHSTVYMGQSSSGGSIHNLQEKGD 1382  
 Db 1302 VQAEYRDAYREYIAQMSQLEGPGSTTISGRSSPHSTVYMGQSSSGGSIHNLQEKGD 1361  
 QY 1383 SEPKDDGRKSFMLKRGDVIDYSSSGVSTNDASPLDPTIETEDEKSDQSGSKLLPGKKSSE 1442  
 Db 1362 SELKQERGRKSFMLKRGDVIDYSSSGVSTNEASPLDPTIETEDEKSDQSGSKLLPGKKSSE 1421  
 QY 1443 RSSLFQTDLKLKSGGLRYQKLPSEDESGTEESDNTPLDKDKRKAEGKVERVPKSPH 1502  
 Db 1422 RPSLFQTDLKLKSGGLRYQKLPSEDESGTEESDNTPLDKDKRKAEGKVERVPKSPH 1481  
 QY 1503 SABPRTFIRAKAYLSDALLDKDSDSGVRSSESSPNHSLHNEVADDSQLKANKLIELE 1562  
 Db 1482 SVEPRTFIRAKAYLSDALLDKDSDSGVRSSESSPNHSLHNEVADDSQLKANKLIELE 1541  
 QY 1563 DDHSGRGRIPHSLSGLQDPIIARMSICSDKSPSCSLTASSPSEENWPACQAYNLNR 1622  
 Db 1542 DEHSGRGRIPHSLSGLQDPIIARMSICSDKSPSCSLTASSPSEENWPACQAYNLNR 1601  
 QY 1623 TPTSIVTLNNSAPANRANQFDEMEGIRETSQVILRFPSSSPNPTTIQENILKSWTHKRSQ 1682  
 Db 1602 TPTSIVTLNNTAPTNRANQFDEIEGIRETSQVILRFPSSSPNPTTIQENILKSWTHKRSQ 1661  
 QY 1683 RSSVTRLSKDPPELHAAASSTGFGGERESIL 1715  
 Db 1662 RSSVTRLSKDPPELHAAASSTGFGGERESIL 1693

RESULT 5

Q80SX9

ID Q80SX9

AC Q80SX9

DT 01-JUN-2003

DT 01-JUN-2003

PRELIMINARY; PRT; 1554 AA.

(TrEMBLrel. 24, Created)

(TrEMBLrel. 24, Last sequence update)





```
Db 1261 TEEDEKSDQSGSKLLPGKSSERPSPSLFQTDLLKKGSLRYOKLPSEDESGTEBSDNTP 1320
Qy 1481 LKODKORKAEGKVERVPKSPHSAEPIRTFIKAEYLSDALDKKSDSGVRSSESPN 1540
Db 1321 LKDDKDKAEGKAERVAKSPHSEVPIRTFIKAEYLSDALDKKSDSGVRSSESPN 1380
Qy 1541 HSLNEVADDSQLEKANILEDDSHSGKRGIPHSLSGLQDPPIARMSCISCDKSPSEC 1600
Db 1381 HSLNEAADDQLEKANILEDEHSGKRGMPHSLSGLQDPVIARMSCISCDKSPSEC 1440
Qy 1601 SLIASSPENPACOKAYNLARTPTSTVTLNNSAPANRANGNFEDEMEGIRTSQVILRPS 1660
Db 1441 SLIASSPEESWPCOKAYNLARTPTSTVTLNNTAPNANQNFDEIEGVRTSQVILRPG 1500
Qy 1661 SSPNPTIQENLKSMTKRSQSRSYTRLSKDPPELHAAASSESTGFGEERESIL 1715
Db 1501 PSPNPTAQVENLKSMAHKRSQSRSYTRLSKDASELH-AASDSTGFGEERESIL 1554

RESULT 6
Q7T163
ID Q7T163 PRELIMINARY; PRT; 1680 AA.
AC Q7T163;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE SI:d2119J18.2 (Novel protein similar to rat kinase D-interacting
DE substance of 220 kDa (KIDINS220) ).
GN Name:ST:d2119J18.2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL954721; CAE17588.1; -.
DR HSSP; P42773; 1IHB.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00023; Ank; 11.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK_REPEAT; 10.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; UNKNOWN_1.
KW ANK repeat; Kinase.
SQ SEQUENCE 1680 AA; 185948 MW; 4F88AB3377DD5E60 CRC64;

Query Match 70.8%; Score 6289.5; DB 2; Length 1680;
Best Local Similarity 72.5%; Pred. No. 2.6e-299;
Matches 1258; Conservative 166; Mismatches 229; Indels 81; Gaps 27;

Qy 1 MSVLISQVINYVEENIPALKALEKCKVDNERNECQOTPLMIAAEOGNLEIVKELIKN 60
Db 9 MTLIAIQNLSFYVEENLAALKVHLDKFEVDGSDNGQTFPLMASQGSLEIVQELIRR 68
Qy 61 GANCNLEDLNWTALISASKEGHVHIVVEELKCGVNLHHRDMGWTALMMAWYKGRTDVV 120
Db 69 GANNVLDVDCWSALISAAKEGHVEVVKELLENSAYIEHRDMGWTALTWASYKGRVEVA 128
Qy 121 ELLLSHGANSVPTGLQVSYVPIIWAAGRHADIIVHLLILONGAKVNCSDKYGTTFPLVWAAR 180
Db 129 TVLLENGANPNNTQQQSVYVPIIWAAGRHAEIVKLELHEGAKVNCSDKYGTTFPLVWAAR 188
Qy 181 KGHLECVKHLAMGADVDQEGANSMTALIVAVKGGYTSQSVKEILKRNPNVNLTKDGNTA 240
Db 189 KGHVDCWHLLENGADVDQEGANSMTALIVAVKGGYTVVVKELKRNPNVNLTKDGNTA 248
Qy 241 LMTIASKEGTEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLQKYADIDIRG 300
Db 1366 EICKLTDKQAEYRNAYEDVIASMSQLELG-----MEKVPVPPFVLSQLMHSSS----- 1366
```

```
Db 249 LMTIAAEGYTEIVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVEIVRALLHKKYADIDIRG 308
Qy 301 QDNKTALYWAKEGNATVRDILQCNPTOTEICTKGETPLIKATQARNIEVVVELLDKGA 360
Db 309 QENKTALYWAKEGNATVRDILQCNPTETTTKDSFPLIKATQARSIEVVVELLDKGA 368
Qy 361 KVSADVKKGDPTPLHIAIRGRSRKLAELLLRNPKDGRLLYRPNKAGETPNIDCSHOKSIL 420
Db 369 KVSADVKKGDPTPLHIAIRGRSRRLAELLLRNPKDGRLLYRPNKAGETPNIDCSHOKSIL 428
Qy 421 TQIFGARHLSHTEGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKSFLLKKL 480
Db 429 TQIFGARHLSHTEGDMGLGYDLYSSALADILSEPTMQPPICVGLYAQWGSKSFLLKKL 488
Qy 481 EDEMKTFAGQOIEPLFQFWSWLI VFLTLLCCGLGLLPFTVHPNLI GAVSLSFALLYIF 540
Db 489 EDEMKTFAGQOIEPLFQFWSWLV LLSLLCCSVALVLGFTVDPKLAIAISILALLYIF 548
Qy 541 FVIYFGGREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPBQTTKALPVRFLF 600
Db 549 FVVYFGSRREGESNNWAWLSTRLARHIGYLELLKLMFVNPPELPBQTTTRALPVRFLF 608
Qy 601 TDYNRLSSVGGETSLAEWIAITLSDACEREFGLATRLFRVFKTEDTQCKKWKTCCLPS 660
Db 609 TDYNRLSSVGGETSMAEMIATLSDACEREFGLATRLFRVFKTEDTQCKKWKTCCLPS 668
Qy 661 FVIFLFIIGCIISGITLAI PRVDPKHLTVNAVLI STASVVG LAFVLCRTWQVLDLIL 720
Db 669 FVIFLFIIGCLIMGNALLAFKVDGQNTVNAVLSMASVVG LALLLCRTWQVTDVIL 728
Qy 721 NSQKRLHNAASKLHLKSEGFMKVLCEVELMARMAKTIDSFTQNTQRLVVIDGLDAC 780
Db 729 NSQKRLHSAANKVHKLSEGFMKVLKNEVELMARMAKTIDGFTQNTQRLVVIDGLDSC 788
Qy 781 EQDKVLOMLDTRVRLPSKGPPIALFASDPHIIIIKAINONLSVLRDSNINGHDYMRNVH 840
Db 789 EQDKVLOMLDTRVRLFSKGPPIISIFASDPHIIIIKAINONLSVLRDSNINGHDYMRNVH 848
Qy 841 LPVFLNSRGLSNARKELVTSATNDVPCSDTGTGQEDADRRVSONSLGEMTKLASKTALN 900
Db 849 LPVFLNSRGLSSAKMKCAPANGET--GNSEGHHEELDKRLQNSLGDQTFKFSKTLN 906
Qy 901 RRDYRERQMQRTITRQMSFDLTLLVTEWDFSDISPTQMRRLNINVSVTGRLLRANQIS 960
Db 907 RRDYRERQMQRSVTRQMSFDLTLLVTEWDFSDISPTQMRRLNINVSVTGRLLRANQIS 966
Qy 961 FNMRLASWINLTQWPYRTSWLILYLEETEGIPDQMTLKIYERISKNIPTTKDVEPLL 1020
Db 967 FNMRLASWINLTQWPYRTSWLILYLEETDGPIDQTNLKIYERISKNIPTTKDVEPLL 1026
Qy 1021 EIDGDIENFEVFLSSRTPVLVARDVKVLPCTVNLDPKRLREIIADVRAAREQISIGLAY 1080
Db 1027 EIDGDIENFEVFLSSRTPVLAARDIRTFPLCTVNLDPKRLREIIADVRAAREQVNMAGTY 1086
Qy 1081 PPLPLEHGPFRAPSGYQPPSVCSST--SFNGPFAAGVVPVSPHSSYYSGMTGPHQPFYNR 1139
Db 1087 PTLPLEQSRP--ISWYSQSSACSPTASFNPNPFGVSPQPHSAYFSGMAGPQHPFYNR 1144
Qy 1140 GSGPAPG--PVLNLSNLDVAVCEKLLKQIEGLDQSMPLPYQYCTTIKKANINGRVLAQCNI 1197
Db 1145 GSASVSGTSPILSSMSTDVCERVKILDGIDQNLISQYTATIKKANINGRVLSQCNI 1204
Qy 1198 ELKKEMMNFGDHLFRSTVLEMRNAESHVVPDPRFLSESSSPAPHEGPAR---ASH 1254
Db 1205 ELKKEMMNFGDHLFRSTVLEMRHVENQVHLSEAP--SEQGSITVGHVEPCRHAGAAQ 1262
Qy 1255 NELPTELSSQTPYTLNFSPEELNTLGLDEGAPRHSNLSWOSQTRRRTPSLSSLSQSSSI 1314
Db 1263 GVAGNTDTPM--YFNFLSFEELS NVGLEB--PPRHVNATWGTTHRTFPSMSSLSQSSSN 1319
Qy 1315 EISKLTDKVQAEYRDAYRIYAQMSQLEGGPGSTTISGRSPSHSTYTYMGSSSGGSIHNS 1374
Db 1320 EICKLTDKQAEYRNAYEDVIASMSQLELG-----MEKVPVPPFVLSQLMHSSS----- 1366
```

QY 1375 LEQEKQSEPKDDGKRSFLMKRG-----DVIDYSSGCVSTNDASPLDPITEDEKSD 1428  
Db 1367 --EDKKKGNDQ--DGRKS--VSRGSGTSGSDNTDYASA-----DAATLDPIITEDEKVD 1416  
QY 1429 QSGSKLLPGKSS-ERSSLFQ--TDLKL--GSLGRYQKLPDEDESGTSESDNTPLKDDK 1485  
Db 1417 HGSKSKLLGRKTSQDKVSLFQADLKLKAGGSRQKLTSDDES--ESDNPFLKDGK 1474  
QY 1486 DRKAEQKVERVPSPEHSAPIRTFTFAKEYLSDALLDKKSDSDGVRSESSPNSHLN 1545  
Db 1475 --KPEAKA-----SDGDRSLTKGDKYLS---DKKSDSDGVRSESSPNSHLQD 1519  
QY 1546 EVADDSQLEKANLIELEDDSHSGKRGIPHSLSGLQDPIIARMSICSDKSPSECSLIAS 1605  
Db 1520 EADLSQSERANLIELEDEANSARKGLPNSLSGLQDPTIARMSICSD-----QCSLLAS 1574  
QY 1606 SPEENPACOKAYNLARTPTVTLNN--NSAPANRANQNFDEMEGIRET--SQVILRPSS 1661  
Db 1575 SPESWPS--SKSYNLARTPTVTLNN--NSAPANRANQNFDEMEGIRET--SQVILRPSS 1633  
QY 1662 SPNPTTONENLKMTHKRSQSRSSYTLKSDKPPELHAAASSESTGFERRESIL 1715  
Db 1634 STTSATTQENVRVHLKRLNGP-----DPPEI-LKVSSTVTFGERRESIL 1680

RESULT 7  
Q6P7Y1  
ID Q6P7Y1 PRELIMINARY; PRT; 1672 AA.  
AC Q6P7Y1;  
DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
DE Hypothetical protein zgc:63531.  
GN ORFNames=zgc:63531;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=Whole body;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Schmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.J., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=Whole body;  
RA Strausberg R.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC061450; AAH61450.1;  
DR HSSP; Q60773; 1AP7.  
DR ZFIN; ZDB-GENE-030131-7824; zgc:63531.  
DR InterPro; IPR002110; ANK.

DR InterPro; IPR001360; Glyco\_hydro\_1.  
DR Pfam; PF00023; Ank; 10.  
DR PRINTS; PR01415; ANKRN.  
DR SMART; SM00248; ANK; 11.  
DR PROSITE; PS0088; ANK REPEAT; 10.  
DR PROSITE; PS0297; ANK REP REGION; 1.  
DR PROSITE; PS00572; GLYCOSYL HYDROL\_F1\_1; UNKNOWN\_1.  
KW ANK repeat; Hypothetical protein.  
SQ SEQUENCE 1672 AA; 185035 MW; 75EBCF10280B410B CRC64;  
  
Query Match 70.6%; Score 6275.5; DB 2; Length 1672;  
Best Local Similarity 72.4%; Pred. No. 1.2e-298;  
Matches 1256; Conservative 166; Mismatches 231; Indels 81; Gaps 27;  
  
QY 1 MSVLSIQSVINVEENIPALKLEKQKVDVERNECGGTPLMIAAEQGNLEIVKELIKN 60  
Db 1 MTTLAQNLFYVEENLAALKVHLDPKFEVDGRSDNGQTPLMLASEQGSLEIVQELIRR 60  
QY 61 GANCNLEDDNNTALISASKEGHVHIVELLKGVNLEHRDMGWTALMWACVKGRTDVV 120  
Db 61 GANVNLDVDCSALLISAAKEGHEVVEVKELLENSAIEHRDMGWTALTWASYKGRVEVA 120  
QY 121 ELLLSHGANSPTGLQYSVYPIIWAAGRHADIHLLQNGAKVNCSDKYGTTPPLVWAAR 180  
Db 121 TVLLEANGNPNTTGGQYSVYPIIWAAGRHAEIVKLLBHGAKVNCSDKYGTTPPLVWAAR 180  
QY 181 KGHLECVKHLAMGADVDOEGANSMTALIVAKGGYTSQSVKEILKRNPNVNLTKDGNTA 240  
Db 181 KGHYDCVHLENGADVDOEGANSMTALIVAKGGYTVVVKELKRNPNVNLTKDGNTA 240  
QY 241 LMTASKEGHEITEVDLLDAGTYVNIIDRSQDVTLVIGAVRGHVEIVRALLQKYADIDIRG 300  
Db 241 LMTAAKEGHEITEVDLLDAGTYVNIIDRSQDVTLVIGAVRGHVEIVRALLHKYADIDIRG 300  
QY 301 QNKATLYWAVEKGNATWVRDILQCNPDTEICTKGTEPLIKATKMRNIEVVELLDKGA 360  
Db 301 QENKATLYWAVEKGNATWVRDILQCNPDTEITTKDSETPLIKATKMRNIEVVELLDKGA 360  
QY 361 KVSADVKKGDTPLHIAIRGSRKLABELLRNPKDGLLRPNKAGETPNYIDCSHOKSIL 420  
Db 361 KVSADVKKGDTPLHIAIRGSRRLABELLRNPKDGLLRPNKAGETPNYIDCSHOKSIL 420  
QY 421 TQIFGARHLSPTETDGMGLYDLYSALADILSEPTMQPPICVGLYAQWGSFGSKFLKKL 480  
Db 421 TQIFGARHLSPTESDGMGLYDLYSALADILSEPTMQPPICVGLNTQWGSFGSKFLKKL 480  
QY 481 EDEMKTFAGQOIEPLFQFSWLIVFLTLCCGGLGLLFAFTVHPNLGIVASLSPLALLYIF 540  
Db 481 EDEMKTFAGQOIEPLFQFSWLVLLSLLCGSVALVGLFTVDPKLAIAISLSLALLYVF 540  
QY 541 FVIVYFGRREGESWNAWVLSLRLARHIGYLELLIKLKFVNPPELPEQTTKALPVRFLE 600  
Db 541 FVIVYFGRREGESWNAWVLSLRLARHIGYLELLIKLKFVNPPELPEQTTKALPVRFLE 600  
QY 601 TDYNRLLSSVGGTSLAEMIALTSACEREFGLATRLFRVFKTEDTQGGKKWKTKCLPS 660  
Db 601 TDYNRLLSSVGGTSLAEMIALTSACEREFGLATRLFRVFKTEDTQGGKKWKTKCLPS 660  
QY 661 FVIFLFIIGCIISGITLLAIFRVDPKHLTVNAVLISIASVVGFLAVLNCRNTWQVLDL 720  
Db 661 FVIFLFIIGCLINGMALLAVFKVQGNQTVNAVLVSMASVVGFLALLNCRNTWQVLDL 720  
QY 721 NSQKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDAC 780  
Db 721 NSQKRLHNAASKLHKLKSEGFMKVLKCEVELMARMAKTIIDSTQNTQRLVLIIDGLDAC 780  
QY 781 EODKVLQMLDTRVRLFSKGPFTAI PASDPHIIKAINQNLNSVLRDSNTHGYDMRNIVH 840  
Db 781 EQDKVLQMLDTRVRLFSKGPFTAI PASDPHIIKAINQNLNSVLRDSNTHGYDMRNIVH 840  
QY 841 LPVFLNSRGLSNARKFLVTSATNGDVPDCSDTTGIGQEDADRRVSONSLGEMTKLGSKTALN 900  
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPDCSDTTGIGQEDADRRVSONSLGEMTKLGSKTALN 900

```
QY 901 RDTYRRRQMTIITRQMSFDLTKLVLTEDWFSDISPQTRMLLNIVSVTGRLLFRANQIS 960
Db 899 RDTYRRRQMSVTRQMSFDLTKLVLTEDWFSDISPQTRMLLNIVSVTGRLLFRANQIS 958
QY 961 FNWDLASWINLQTEOMPRTSWLILYLBETEGIPQMTLKIYERISKNITTKDVEPLL 1020
Db 959 FNWDLASWINLQTEOMPRTSWLILYLBETEGIPQMTLKIYERISKNITTKDVEPLL 1018
QY 1021 EIDGDIRNFVPLSRTPLVARDVKVFLPCTVNLDPKRLREIADVRAAREQISIGGLAY 1080
Db 1019 EIDGDIRNFVPLSRTPLVARDVKVFLPCTVNLDPKRLREIADVRAAREQISIGGLAY 1078
QY 1081 PPLPLHEGPRAPSGYSPVCSST-SFNGPFPAGGVVSPPHSSYYSGMTGPQHPFFYR 1139
Db 1079 PPLPLHEGPRAPSGYSPVCSST-SFNGPFPAGGVVSPPHSSYYSGMTGPQHPFFYR 1136
QY 1140 GSGPAPG--PVVLLNSLVADVCEKLOIEGLDQSMPLFOYCTTIKANINGRVLAQCNID 1197
Db 1137 GSGPAPG--PVVLLNSLVADVCEKLOIEGLDQSMPLFOYCTTIKANINGRVLAQCNID 1196
QY 1198 ELKEMNNFVGDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARR---ASH 1254
Db 1197 ELKEMNNFVGDWHLFRSTVLEMRNAESHVVPEDPRFLSESSSGPAPGEPARR---ASH 1254
QY 1255 NELPHELSQTPYTLNFSFEELNLTGLDEGAPRHSNLSWOSQTRRTPLSLSLNSQDSSI 1314
Db 1255 GVGWNTDTPM--YFNFLSPEELSNVLEB-PPRHVNATWGCATHTFMSLSLNSQESSN 1311
QY 1315 EISKLTDKQBYRDAYREYIAQMSQLGEGGPGSTTISGRSPHSPTTYNGQSSGGSISHN 1374
Db 1312 EICKLTDKQBYRDAYREYIAQMSQLGEGGPGSTTISGRSPHSPTTYNGQSSGGSISHN 1374
QY 1375 LEQKGGKDEPKDPDGRKSLFKMG- ----DVIDYSSGVSSTNDASPLDPTREDEKSD 1428
Db 1359 --EUKKGDGNDQ--DGRKS-VSKRGSTKSGDNTDYASA----DAATLDPITBEDEKVD 1408
QY 1429 QSGSKLLPGKISS-ERSLFO--TDLKLK-GSGLRYQKLPSEDESGETESDNTPLLKDDK 1485
Db 1409 HGSSKSLGRKTSKGLKVSFLQADLKLKAGGSRVQKLTSDDES--EESDAPLKKDK 1466
QY 1486 DRKAEKVERVPKPEHSAEPIRTIKAKYLSDALDKDSDSGVRSSESSPNHSLN 1545
Db 1467 --KPEAKA-----SDGGDRSLTKGKVLK----DKDSDSGVRSSESSPNHSLQD 1511
QY 1546 EVADDSOLEKANLIELEDDSHSGKGIPIHSLSLQDPTIARMSICSEDKKSPSECSLIAS 1605
Db 1512 BEADLSOSERANLIELEDDSHSGKGIPIHSLSLQDPTIARMSICSEDKKSPSECSLIAS 1566
QY 1606 SPEENWPACQAKYNNLRTPSTVTLNN--NSAPANRANQNFDEMEGIRET--SQVILRPS 1661
Db 1567 SPEENWPACQAKYNNLRTPSTVTLNN--NSAPANRANQNFDEMEGIRET--SQVILRPS 1625
QY 1662 SPNPTTIONELKGMTHKRSORSSYTRLSKDPPELHAAASSESTGFGERESEIL 1715
Db 1626 STTSATTQNNVVRVHLKRLNPG-----DPPEI-LKVSSETVTFGERESEIL 1672
```

## RESULT 8

```
Q9UF42 ID Q9UF42 PRELIMINARY; PRT; 1031 AA.
AC DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp434F0621.
GN Name=DKFZp434F0621;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE-Testis;
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133620; CAB63746.1; -.
DR PIR; T43458; T43458.
DR HSSP; P09959; ISW6.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank, 11.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK, 11.
DR PROSITE; PS00088; ANK_REPEAT; 10.
DR PROSITE; PS02097; ANK_REPEAT; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 1031 AA; 115310 MW; D98866461C13A2F5 CRC64;

Query Match 58.6%; Score 5205.5; DB 2; Length 1031;
Best Local Similarity 98.5%; Pred. No. 1.6e-246;
Matches 1008; Conservative 2; Mismatches 0; Indels 13; Gaps 1;

QY 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECCQTPLMIAAEEQNLKELIKN 60
Db 1 MSVLISQSVINYVEENIPALKALLEKCKVDNERNECCQTPLMIAAEEQNLKELIKN 60
QY 61 GANCNLELDNWTALISASKEGHVHVEELLKCGVNLKCGVNLKCGVNLKCGVNLKCGVNL 120
Db 61 GANCNLELDNWTALISASKEGHVHVEELLKCGVNLKCGVNLKCGVNLKCGVNLKCGVNL 120
QY 121 ELLLSHGANSVTCGLQSVYPIIWAAGRGHADIHLLHLLONGAKVNCSDKYTTPLVWAAR 180
Db 121 ELLLSHGANSVTCGLQSVYPIIWAAGRGHADIHLLHLLONGAKVNCSDKYTTPLVWAAR 180
QY 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
Db 181 KGHLECVKHLAMGADVDEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDGNNTA 240
QY 241 LMTASKEGTEIVODLLDAGTYNNIPDRSGDTVLIGAVRGGHVEIVRALLQYADIDIRG 300
Db 241 LMTASKEGTEIVODLLDAGTYNNIPDRSGDTVLIGAVRGGHVEIVRALLQYADIDIRG 300
QY 301 QDNKTALYNAVEKGNATVRDILQCNPDTEICTKGDTPLIKATMRNIEVEVLLDKGA 360
Db 301 QDNKTALYNAVEKGNATVRDILQCNPDTEICTKGDTPLIKATMRNIEVEVLLDKGA 360
QY 361 KVSADVKKDTPHIAIRGRSRKLAELLLRNPKDGLLYRPNKAGETPYNIDCSHOKSIL 420
Db 361 KVSADVKKDTPHIAIRGRSRKLAELLLRNPKDGLLYRPNKAGETPYNIDCSHOKSIL 420
QY 421 TQIFGARHLSPTEITDGMGLYDLYSSALADILSEPTMQPPTCVGLYQWGSFGKFLKKL 480
Db 421 TQIFGARHLSPTEITDGMGLYDLYSSALADILSEPTMQPPTCVGLYQWGSFGKFLKKL 480
QY 481 EDEKMTFAGQIEPLFQPSWLIIVFLTLILCGGLGLLPAFTVHPNLGIAVSLFLALYIF 540
Db 481 EDEKMTFAGQIEPLFQPSWLIIVFLTLILCGGLGLLPAFTVHPNLGIAVSLFLALYIF 540
QY 541 FVIYFGRRREGESNNWAVLSTRARHIGVLELLKLMFVNPPPELQTTKALPVRLFL 600
Db 541 FVIYFGRRREGESNNWAVLSTRARHIGVLELLKLMFVNPPPELQTTKALPVRLFL 600
QY 601 TDYNRLSSVSGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKKKTKCLIPS 660
Db 601 TDYNRLSSVSGGETSLAEMIATLSDACEREFGLATRLFRVFKTEDTQKKKKKTKCLIPS 660
QY 661 FVIFLFIIGCIITGIIITLAIIPRVDPKHLTVNAVLIASVVGGLAVFLNCRTWQVLSLL 720
Db 661 FVIFLFIIGCIITGIIITLAIIPRVDPKHLTVNAVLIASVVGGLAVFLNCRTWQVLSLL 720
QY 721 NSQKRLHNAASKLHLKLSGFMVKLCEVELMARMARKTIDSFONOTRLVVIIDGLDAC 780
Db 721 NSQKRLHNAASKLHLKLSGFMVKLCEVELMARMARKTIDSFONOTRLVVIIDGLDAC 780
QY 781 EQDKVLQMLDTRVRLFSKGPFFIAIFASDPHIIIKAINQNLNLSVLRDSNINSHDYMRIHVH 840
Db 781 EQDKVLQMLDTRVRLFSKGPFFIAIFASDPHIIIKAINQNLNLSVLRDSNINSHDYMRIHVH 840
```

```
Db 781 EQDKVLQMLDTRVLFSGKPFIAFAFDPHIIKAINQNLNSVLRDSNINGHDYMRNIHV 840
Qy 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCSTTGIQEDADRRVSONSLGEMTKLGSXTALN 900
Db 841 LPVFLNSRGLSNARKFLVTSATNGDVPSCSTTGIQEDADRRVSONSLGEMTKLGSXTALN 900
Qy 901 RRTYTRRRQWRTITRQMSFDLTKLAVTEWDFDISPQTMRRLLNVSVTGRLLRANQIS 960
Db 901 RRTYTRRRQWRTITRQMSFDLTKLAVTEWDFDISPQTMRRLLNVSVTGRLLRANQIS 960
Qy 961 FNDRLASWINLTQMPYRTSWILYLLEETEGIPDQMTLKIYER-----IS 1007
Db 961 FNDRLASWINLTQMPYRTSWILYLLEETEGIPDQMTLKIYERCCGADSCDRDRIGIS 1020
Qy 1008 KNI 1010
Db 1021 KSV 1023

RESULT 9
Q72322 PRELIMINARY; PRT; 691 AA.
AC Q72322;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZp686H14204;
GN Name=DKFZp686H14204;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Human fetal brain;
RC Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX58189; CAD98059.1; -;
DR InterPro; IPR010993; SAM_homology.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 691 AA; 75760 MW; F5783F24DDEA6027 CRC64;

Query Match 40.3%; Score 3583; DB 2; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.9e-167;
Matches 691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1025 DIRNEFVLSRTPVLVARVKVFLPCTVNLDPKLRRIADVRAAREQISIGGLAYPLP 1084
Db 1 DIRNEFVLSRTPVLVARVKVFLPCTVNLDPKLRRIADVRAAREQISIGGLAYPLP 60

1085 LHGGPPRAPSQSGYQPPSVCSSTSFNGPFGAGVVSPPHSSYSGMTGPQHPFYNRGSGPA 1144
Db 61 LHGGPPRAPSQSGYQPPSVCSSTSFNGPFGAGVVSPPHSSYSGMTGPQHPFYNRGSGPA 120

1145 PGVVLNSLVNDAVCEKLRQIEGLDQSMLPQYCTTIKKANINGRVLACNIDELKEMN 1204
Db 121 PGVVLNSLVNDAVCEKLRQIEGLDQSMLPQYCTTIKKANINGRVLACNIDELKEMN 180

1205 MNFGDWHLFSTVLEMRNAESHVVPDPFLSSSSSGPAPHGPARRASHNELPHTLELS 1264
Db 181 MNFGDWHLFSTVLEMRNAESHVVPDPFLSSSSSGPAPHGPARRASHNELPHTLELS 240

1265 QTYVTNLNFSPEELNTGLDGCAPRHNLNMQSOTRTRTPSLSSNSODSSIEISKLTDKQV 1324
Db 241 QTYVTNLNFSPEELNTGLDGCAPRHNLNMQSOTRTRTPSLSSNSODSSIEISKLTDKQV 300

1325 AEYRDAYREYIAQMSOLEGGPGSTTISGRSSPHSTHYMGQSSGGSIHNLBEQEKGDSE 1384
Db 301 AEYRDAYREYIAQMSOLEGGPGSTTISGRSSPHSTHYMGQSSGGSIHNLBEQEKGDSE 360

1385 PKDDGKSLMKRGDVIDYSSSGVSNDSPLDPIITEDEKSDQSGSKLLPGKXSERS 1444
```

```
Db 361 PRPDDGRKSLMKRGDVIDYSSSGVSTNDASPLDPIITEDEKSDQSGSKLLPGKXSERS 420
Qy 1445 SLFQTDLKLKSGRLYQKLPSEDEDSGTESDNTPLLKDDKDKAKCKVERVPKSPHSA 1504
Db 421 SLFQTDLKLKSGRLYQKLPSEDEDSGTESDNTPLLKDDKDKAKCKVERVPKSPHSA 480
Qy 1505 EPRTPIKAKYKLVSDALLDKDSSDGVSRSSSPNHSLNHVADDSOLEKANLIELEDD 1564
Db 481 EPRTPIKAKYKLVSDALLDKDSSDGVSRSSSPNHSLNHVADDSOLEKANLIELEDD 540
Qy 1565 SHSGKRGIPHSLGSLQDPIIARMSICSEDKSPSECSLIASPEENWPACQKAYNLNRT 1624
Db 541 SHSGKRGIPHSLGSLQDPIIARMSICSEDKSPSECSLIASPEENWPACQKAYNLNRT 600
Qy 1625 STVTLLNNSAPANRANQNFDEMEGIRETSQVILRPSSSNPPTTIQENILKSMTHKSORS 1684
Db 601 STVTLLNNSAPANRANQNFDEMEGIRETSQVILRPSSSNPPTTIQENILKSMTHKSORS 660
Qy 1685 SYTRLSKOPPELHAAASSESTGGERESIL 1715
Db 661 SYTRLSKOPPELHAAASSESTGGERESIL 691

RESULT 10
Q6MZU2 PRELIMINARY; PRT; 664 AA.
ID Q6MZU2;
AC Q6MZU2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686A19189 (Fragment).
GN Name=DKFZp686A19189;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=FROM N.A.
RC TISSUE=Human esophagus tumor;
RG The German Human cDNA Consortium;
RA Bloecker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640878; CAE45935.1; -;
DR InterPro; IPR010993; SAM_homology.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 664 AA; 72879 MW; 678BA3EF27DA1BDF CRC64;

Query Match 36.1%; Score 3211; DB 2; Length 664;
Best Local Similarity 94.0%; Pred. No. 4.7e-149;
Matches 624; Conservative 0; Mismatches 2; Indels 38; Gaps 1;

Qy 1090 PRAPSGYQPPSVCSSTSFNGPFGAGVVSPPHSSYSGMTGPQHPFYNR----- 1139
Db 1 PRAPSGYQPPSVCSSTSFNGPFGAGVVSPPHSSYSGMTGPQHPFYNRFPAPLYTP 60
Qy 1140 -----GSGPAGPFWLLNSLVNDAVCEKLRQIEGLDQ 1171
Db 61 RYYPGSGQHLISRPSVKTSIPLRQNNGLSGPAGPFWLLNSLVNDAVCEKLRQIEGLDQ 120

1172 SMLPQYCTTIKKANINGRVLACNIDELKEMNFGDWHLFSTVLEMRNAESHVVPED 1231
Db 121 SMLPQYCTTIKKANINGRVLACNIDELKEMNFGDWHLFSTVLEMRNAESHVVPED 180

1232 PRFLSSSSSGPAPHGPARRASHNELPHTLELSQTYTNLNFSEELNTGLDGCAPRHNS 1291
Db 181 PRFLSSSSSGPAPHGPARRASHNELPHTLELSQTYTNLNFSEELNTGLDGCAPRHNS 240

1292 LSWQSOTRTRTPSLSSNSODSSIEISKLTDKQAEYRDAYREYIAQMSOLEGGPGSTTIS 1351
Db 241 LSWQSOTRTRTPSLSSNSODSSIEISKLTDKQAEYRDAYREYIAQMSOLEGGPGSTTIS 300
```

QY 1352 GRSSPHSTYMGQSSSGSIHNSLQKKGKSEPKDGRKSFMLMKRGVDIYSSSGYST 1411  
 DB 301 GRSSPHSTYMGQSSSGSIHNSLQKKGKSEPKDGRKSFMLMKRGVDIYSSSGYST 360  
 QY 1412 NDASPLDPTREDEKSDGSKLLPGKSSSRSLFOTDLKLGSLRYQKLPSEDESG 1471  
 DB 361 NDASPLDPTREDEKSDGSKLLPGKSSSRSLFOTDLKLGSLRYQKLPSEDESG 420  
 QY 1472 TEESDNTPLKDDKDRKAEKGVKPSPEHSAEPIRTFIKAEKYLSDALLDKKSSDSG 1531  
 DB 421 TEESDNTPLKDDKDRKAEKGVKPSPEHSAEPIRTFIKAEKYLSDALLDKKSSDSG 480  
 QY 1532 VRSSSSPNHSLHNEVADDSOLEKANLIELEDDSHSGRGKIPHSLSGLQDPIIARMSTCS 1591  
 DB 481 VRSSSSPNHSLHNEVADDSHLEKANLIELEDDSHSGRGKIPHSLSGLQDPIIARMSTCS 540  
 QY 1592 EDKSPSCSLIASSPEENWAPACQAYNLNRTPTVTTLNNSAPANRANQNFDEMEGIRE 1651  
 DB 541 EDKSPSCSLIASSPEENWAPACQAYNLNRTPTVTTLNNSAPANRANQNFDEMEGIRE 600  
 QY 1652 TSQVILRPSSSPNPTTIQENLKSMTKRSORSSYTRLSKDPPELHAAASSTGFGGER 1711  
 DB 601 TSQVILRPSSSPNPTTIQENLKSMTKRSORSSYTRLSKDPPELHAAASSTGFGGER 1711  
 QY 1712 ESIL 1715  
 DB 661 ESIL 664

RESULT 11  
 Q9H9E4  
 ID Q9H9E4 PRELIMINARY; PRT; 543 AA.  
 AC Q9H9E4;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hypothetical protein FLJ12811.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,  
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
 RA Tanai H., Kinata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,  
 RA Ono Y., Takiguchi S., Watanabe S., Yoseida M., Hotuta T., Kusano J.,  
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,  
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
 RA Musashino K., Yuuki H., Oshima N., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,  
 RA Moriya S., Momiya M., Sato N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yanazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yanashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;

"Complete sequencing and characterization of 21,243 full-length human  
 cDNAs.";  
 RT Nat. Genet. 36:40-45(2004).  
 RL EMBL; AK022873; BAB14285.1; -  
 DR InterPro; IPR010993; SAM homology.  
 SQ SEQUENCE 543 AA; 59976 MW; FSD643D5A20C641D CRC64;  
 Query Match 31.4%; Score 2791; DB 2; Length 543;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-128;  
 Matches 541; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1173 MLPOYCTTIKANKINGRVLAQCNDIDELKKNMNFQDWHLFRSTVLEWRNAESHVVPD 1232  
 DB 1 MLPOYCTTIKANKINGRVLAQCNDIDELKKNMNFQDWHLFRSTVLEWRNAESHVVPD 60  
 QY 1233 RFLSESSGPPHGPARRASHNELPHELSSQTPYTLNFSFEELTGLDEGAPRHNL 1292  
 DB 61 RFLSESSGPPHGPARRASHNELPHELSSQTPYTLNFSFEELTGLDEGAPRHNL 120  
 QY 1293 SWQQTTRTPSLSLNSQDSSIEISKLTDKQAEYRDAYREYIAQMSQLEGPGSTTISG 1352  
 DB 121 SWQQTTRTPSLSLNSQDSSIEISKLTDKQAEYRDAYREYIAQMSQLEGPGSTTISG 180  
 QY 1353 RSSPHSTYMGQSSSGSIHNSLQKKGKSEPKDGRKSFMLMKRGVDIYSSSGYSTN 1412  
 DB 181 RSSPHSTYMGQSSSGSIHNSLQKKGKSEPKDGRKSFMLMKRGVDIYSSSGYSTN 240  
 QY 1413 DASPLDPTREDEKSDGSKLLPGKSSSRSLFOTDLKLGSLRYQKLPSEDESGT 1472  
 DB 241 DASPLDPTREDEKSDGSKLLPGKSSSRSLFOTDLKLGSLRYQKLPSEDESGT 300  
 QY 1473 EESDNTPLKDDKDRKAEKGVKPSPEHSAEPIRTFIKAEKYLSDALLDKKSSDSGV 1532  
 DB 301 EESDNTPLKDDKDRKAEKGVKPSPEHSAEPIRTFIKAEKYLSDALLDKKSSDSGV 360  
 QY 1533 RSSESSPNHSLHNEVADDSOLEKANLIELEDDSHSGRGKIPHSLSGLQDPIIARMSTCS 1592  
 DB 361 RSSESSPNHSLHNEVADDSHLEKANLIELEDDSHSGRGKIPHSLSGLQDPIIARMSTCS 420  
 QY 1593 DKKSPSCSLIASSPEENWAPACQAYNLNRTPTVTTLNNSAPANRANQNFDEMEGIRET 1652  
 DB 421 DKKSPSCSLIASSPEENWAPACQAYNLNRTPTVTTLNNSAPANRANQNFDEMEGIRET 480  
 QY 1653 SQVILRPSSSPNPTTIQENLKSMTKRSORSSYTRLSKDPPELHAAASSTGFGGER 1712  
 DB 481 SQVILRPSSSPNPTTIQENLKSMTKRSORSSYTRLSKDPPELHAAASSTGFGGER 540  
 QY 1713 SIL 1715  
 DB 541 SIL 543

RESULT 12  
 Q7KVP5  
 ID Q7KVP5 PRELIMINARY; PRT; 1604 AA.  
 AC Q7KVP5;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE CG30387-PC.  
 GN ORFNames=CG30387;  
 DN Drosophila melanogaster (Fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Spidroidae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,









OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachyoptera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazew R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA April J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong Y., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskaas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 RT *melanogaster* euchronomic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0073-RESEARCH0079 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whittied E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003453; AAF46711.4; -.  
 DR HSPF; P42773; 11HB.  
 DR INACT; Q9W210; -.  
 DR InterPro; IPR002110; ANK.  
 DR Pfam; PF00023; Ank; 10.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 12.  
 DR PROSITE; PS50088; ANK\_REPEAT; 9.  
 DR PROSITE; PS50297; ANK\_REPEAT; 9.  
 KW ANK repeat.  
 SQ SEQUENCE 1678 AA; 184516 MW; D75A23F12B545DFC CRC64;  
 Query Match 27.2%; Score 2417; DB 2; Length 1678;  
 Best Local Similarity 35.8%; Pred. No. 1.5e-103;  
 Matches 597; Conservative 286; Mismatches 552; Indels 232; Gaps 48;  
 QY 1 MSVLISQVINYVEENIPALKALLE-KCKDVERNECGQTPLMTAEEQGNLEIYKELIK 59  
 DB 75 MSLGHRALLQYIDNNDDLSGLRLTDDRDENATTVLVVAVRGRTAFVREFLA 134  
 QY 60 NGANCNLEDLNTALISASKEGHVHIVEELLKCGVNLHEDMGWGTALMWACYGRDIDV 119  
 DB 135 RGADVQAEEDLNTALLCASRNGHLDVVQLLDHGAEEVHRDGGWTSLWAAVGHTEL 194  
 QY 120 VELLISHGANSVTGLQSVVPIIWAAGRHADIVHLLONGAKVNCSDKYGTTPVLWAA 179  
 DB 195 VRLLLDKGADGNAHG-NYHLGALLWAAAGRYKDIVELLVQRGAKVNVGDKYGTTLVWAC 253  
 QY 180 RKGHLECKVHLLANGADVQEGANSMTALIVAVKGYTQSVKEILKRNPNVNLTKDQNT 239  
 DB 254 RRGVNEIVDTLLKAGANVDTAGMYSWTPLLVAAGGGHTDCVSSILEKKPNVNLDDQMT 313  
 QY 240 ALMTASKEGHTPIVQDLIDAGTYVNIPIRSQDVTVLIGAVRGHVEIVRALLQKVIADIR 299  
 DB 314 ALCTASREGFDIAASLIAGAYINIQRGADTPLIHAVKAGHRTVVEALLKGHADVDIQ 373  
 QY 300 QDNKTALYWAWEKGNATMVRDIILOCPNPTDEICTQGETPLIKATPMKNIEVVELLDKG 359  
 DB 374 GKDRKTAIYAVEKGHPTIVKLLLATNPDLSESATKDGDTPLLRVNRNLEIVHLLDRK 433  
 QY 360 AKVSADVCKGDTPLHAIIRGSRKLAELLNPKDGLLYRPNKAGEPYINDCSHKSI 419  
 DB 434 AKVTASDKRGDTCLHIAWRAKSKTIVKALLNKHSLLYRANKAGETPYINLDSHQTI 493  
 QY 420 LTQIFGARHLSPTDGDMLGYDLYSSALADILSEPTMOPPICVGLVAQWGSKGSFLKK 479  
 DB 494 LGQVFGARRLNTNEDSEGMLGYELYSALADVLSEPTLTITTYTGLYAKWGSKGSFLNK 553  
 QY 480 LEDEMKTFAGQOIPLPQFQSWLIVFLTL---LLCG---GLGLLFAFTVHPNLGIAVSUSF 533  
 DB 554 LRDEMNNFARQWASPPRTSTGLLFIIVCLHVALLIGTIVGLTSWASV-----VGVSAAVGF 608  
 QY 534 LALLYIFPIYVIFGRRREGESWNWVLSFLARHIGYLELLLLKLMFVNPPELPEQTTKA 593  
 DB 609 LLLAYLLAAVRYCNYQ--MDMOWAYSQVHGKRMTRLRLLIQVAFCHPFG--PQSDQSA 665  
 QY 594 LPVRFLFTDYNRLSSVGSETSIAEWIATLSACREFGLATRLFRVFKTE--DTQGGKK 651

```
Db 666 KPRFHAENASAPT-GDCAVHMLAALDALSHYGLWATLYRAFRPKCLKVDVGWR 724
QY 652 WKTCCLPSVIFLFIIGIISGITLAI- RVDPKHLTVNAVLSIASVWGLAFV 706
Db 725 WRMCCIPVILFELALVTVTGISTVAVFTFADEKEKEHILV-ALYVIAAVMGTLIC 782
QY 707 LNCRTWVQLDLSLNSQKELHNAASKLHKLSEGFMKVLKCVELMARMAKTIDSTQN 766
Db 783 THLVLAKVEVSFTSHIRVLKAV- RSSSAPLTMIGAEVAVWMTVMKCLDAPTNQ 838
QY 767 QTRLVIIIDGLDACEQKVQLMDLTVRVLFS- KGPPIAIFASDPDPIIIRKAINQNSVL 824
Db 839 QSLVGVIDALDSCDTERIILTNVQTLSSPNRPVLLISVDPHVIAKAAENGRRLF 898
QY 825 RDSNINGHYMNVHLVFLNSRGLSNARKFLVTS- ANGVDPVCS- TTGIQED 877
Db 899 TEGGICGHDFLRNLVHLVPLVQLNSGLRKVQRAQMTALLPKRSGGDYQTDGPTLGHVS 958
QY 878 ADRVQSNS- LGEMTKL- GSKTALNR- RTYRRQMQRTITRQMSF 920
Db 959 A-RRLSNASEIISQEKLRGAPRGGGKRLRLSESVASSTGSLNHLRQNPQTV- L 1012
QY 921 DLTKLAVTEWDFDISPOTVRRLLINIVSVTGRLLRANOISFNWDRLASWINLTEQWPYRT 980
Db 1013 DLSRVLTDYDFDVNPRSNRMLNVIYIIVLLKAFQEFYSWYRLSSWINLTEQWPLRA 1072
QY 981 SMILYLEE- TEGIPQWMTKTYIRISKNIPPTKDVPELLEIDGDIRNFVFLSSRTP 1038
Db 1073 SMIVLHHDQPMDSNADESLSQSVYEKRLPKLAYREAAPLLELDRDKLDAFLQLHKS 1132
QY 1039 VLVARVVKVLPCTVNLDPKRLIAD- VPAAREQISIGGLAVPPLPHE 1087
Db 1133 DLLVADLRIFLPTIINLDPYLRKVEDQOTIEDEGLSVIQAARPSVNTWRQP- 1186
QY 1088 GPPRAPSGVQSPVSCVSSTFNGPFAGVVPVPPHSSY- YSGMTG 1131
Db 1187 -AFTY- VSPQAPYPPYQFONEYPANELRSNLSTSTE 1223
QY 1132 PQHPFYNRGSGPAPV- LNLNSLVADYCEKLQIEGLDSMLPOYCTTIKANINGR 1189
Db 1224 PVTPLNSPDSGDIIQKTLDTAVGVISLDRIDM- KPALPKLAPVLRENAINGR 1282
QY 1190 VLAQCNIDELKEMWNGFWHLFRSTVLEMRNAESHVP- EDRPFLSESSGP 1242
Db 1283 VLKHCMPDLKSLVLSFGHWELFRLLITTLRECER- LPRKQOQOQPGALEAPSNVP 1340
QY 1243 APHG- EPARR- ASHNELPHE- LSSQ- TPYTLNFSFELNLTGLDEGA 1286
Db 1341 MIKDVTDALMQPPRESLSRKNSVSHMEKQVTLSEQMLCGTLQTLNEAYE- DVASSERPS 1399
QY 1287 PRHSNLSWQSQRTRTPSLSL- NSQDSSIEISKLTDKV- AFY-RDAYREYI 1335
Db 1400 PTCEMLAAVAQLQAPLIPRESSEFGSPDDQKQYGVKISNNNNNNQYLHAEYNSVSSHLS 1459
QY 1336 AQMSQLEGPGSTISGRSPH- STY- YMGQSSSG- SI 1371
Db 1460 QSLSTLVGAPVHGGSGGSHLNGNDLSDSTLDLMHVDSVFGGGGGGYHRSRQISI 1519
QY 1372 HSNLEQKGD- SPKPDGGRKSPMKRKGDV- IDYS 1406
Db 1520 SSELLHESKLDTPSAGAMTTPLLGASGVAP- SGRESLLKQOGSVKADKRVSIQOMAT 1578
QY 1407 SGVSTNDASLDP- ITEDEKSPQSGSKLL- PGKK- SSERSLFTQDL 1451
Db 1579 SNNNNNNKLTJFNVEVVSERQEVQAGKRLTTKPPGPRPASLIITRNDNSQFQL- 1636
QY 1452 KLKSGSLRYQLRSDDESGTESDNTPLKDDKDKRKAEGKVERVPK 1498
Db 1637 -LRSSSVYDDVEAQEHRT- TIRTTLEQEEBESAPFVFTVRK 1678
```

RESULT 15

```
Q7PZY2
ID 07PZY2 PRELIMINARY; PRT; 1354 AA.
AC Q7PZY2;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Ebip9166 (Fragment).
GN Name=ebiG9166; ORNames=ENSANGG00000006867;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data
CC EMBL; AAB01008986; EAA00198.1; -.
DR HSSP; O75832; IUOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 11.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS50088; ANK_REPEAT; 9.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR ANK repeat.
FT NON TER 1354 1354
FT NON TER 1354 1354
SQ SEQUENCE 1354 AA; 148250 MW; E0814F869C911280 CRC64;

Query Match 26.2%; Score 2328.5; DB 2; Length 1354;
Best Local Similarity 39.0%; Pred. No. 2.4e-105;
Matches 525; Conservative 243; Mismatches 440; Indels 137; Gaps 31;

QY 1 MSVLISQSVNVYVERENIPALKALL-EKCDVDNERCEGOTPLMIAEQNLISVIELIK 59
Db 3 MSLGHSRLQYLETDDDLGSLKSLFGLTRHLQVDRDNNVTVLVWASGRGATHFVELLA 62
QY 60 NGANCNLELDNWNLTALISKEGHVHIVEELLKCGVNLHRDMGWTALMWACYKGRTDV 119
Db 63 RGADVQAQDLDSWTALHFAAKAGHVGIVELLNDNGAELEHRDMGWTALMWGSKGHTSV 122
QY 120 VELLISHGANPSVTGLQSVYPIIAAGRGHADIIVHLLIQ-NGAKVNCSDKYITPLVWA 178
Db 123 VALLLQRGADVQAHG-NYHLNPLLWASGRGHTIEVLLLVNTGKGVNVDGKYITPLVWA 181
QY 179 ARKGHLECKVHLLAMGADVDOEGANSMTALIVAVKGGYTQSVKELKRNPNVNLTKDGN 238
Db 182 CRKGSREIVDVLKAGANVDTAGMYSWTPLLVAVSGGFQECVSLLLERKPNVNLKXGM 241
QY 239 TALMIASKEGHTIEVODLLDAGTYVNIIPDRSGDTVLIGAVRGGHVIRVALLQYADIDI 298
Db 242 TALSIAACREGLTEIASALIAAGAYLVNQDRAGDTPLINAVKGGHRSVVEILMKRHVDVI 301
QY 299 RGDNNTALYAVKGNATWVRDILOCNPDTEICTDGETPLIKATKMRNIEVVELLDK 358
Db 302 QGKDKKTALYTAVEKGHTAIVKILQSNPDLELSTKDGDTALLRAVRNRLVQMLER 361
QY 359 GAKVSAVDKGGDTPLHIAIRGRSRKLAELLRNPKDRLLYRPNKAGETPNYDCSHOKS 418
Db 362 KAKVGATDKRGDTCLHVAWRARSKAIVEALLSNPKYQLLYRSNKGETPYALDTHQKT 421
QY 419 ILTQIFGARHLSPTETDGMGLGYDLYSSALADILSEPTMQPPTCVGLYAGWGSKGFLLK 478
Db 422 ILQVFGNRLNANEDSEGMLGYLYSSALADVLSEPTLTPTITVGLYAKWGSKGFLLT 481
QY 479 KLEDEKMTAGQOIEPLFOFSWLVIVLTLCCGLGLLFAFTVPHNL-GIANSLSFLALL 537
Db 482 ELRDEKMFNAHSWSEPPIDASLFFLIISLFLVIVGLATVSWVGWMTGATILLVI 541
QY 538 YIFFIVYFGRRREGSSWNWVLSLRLARHIGYLELLKLMFVNPPPEQTTKALPVR 597
```

Db 542 YFTDILLKPLDR--YDLEWYSLNYGLSKLGRRLILQVAFCHPPG-POSQOQMEVR 598  
Qy 598 FLFTDYNRLSSVGGSTSLAEMIATLSDACERFEGFLATRLFRVFKTEDTQ--GKKKWKKT 655  
Db 599 FHFAEASG-AAPNGDAVALMLASLFDALIAHYGSLATGLYRAFPPKPKATGGWKWKRM 657  
Qy 656 CCLPSFVFLFIIGCIISGILLAIIFRVPDKHLTVNA--VLISIASVVGAP---VLNC 709  
Db 658 CCMP--VVLMEFELGML--GILATASLSIVYSEYGLGEBEIAVAIYVLLGILLAGAIANL 713  
Qy 710 RTWQVLDLSLNSQKRLHNAASKHLKLSGFMKVLKCEVELMARMMAKTIDSFTQOTR 769  
Db 714 HAWSKLIGALFMSQGHKLRAPNS----NEAAPTALGAEVSLMTDMVRCCLDAFTGQOSR 769  
Qy 770 LVVIIDGLDACQDKVLOMLDTRVLFSS--KGPFAIFASDPHIIIIKAINONLNSVLKDS 827  
Db 770 LVGVVDALDSCDTERLTILNAVQTLISGFORPPFVILLAVDPHVIKAAEANSRLFTTEG 829  
Qy 828 NINGHDYNRNIVHLPVFLNSRGLSNARKFLVT--SATNGDVPCSDTTGTGIOEDAD----- 879  
Db 830 GIGGHDPLRNLVHLPVYLQNSGLKRVQRAQNTAMSAVRRAMP-----DLSRDDDDQPHLGH 884  
Qy 880 ----RRYQONS--LGEWTKLGSKTALNRDTRRRMQ-----RTITRQMSFDLT 923  
Db 885 SASVRLRLSNASEIMSSQEKLRAMPSSRSGSKLRVSDSIASSIGSNLHKLQNPVVDLS 944  
Qy 924 KLLVTEDWFDISPOWRLRLNIVSVTCRLLRANQISFNWDLASWINLTEOWPVRTSWL 983  
Db 945 KILITDDYFSDVNRMRRLMNVIYITVLLKAFQIDFSWYRLSSWINLTEOWPLRSMI 1004  
Qy 984 ILYLEETEG---IPDQWTKTIYERISKNIPTTKDVEPLLEIDGDIRNFEVFLSSRTPLV 1040  
Db 1005 V--LEHDQGBDSFDDSVSLQAVDKVRPKLTCLEAAALLDLDRDKLDAFLQLHKSXL 1062  
Qy 1041 VARDVKVLPCTVNLDPKLEIADVRAREQ-----ISIGGLAYPPLPLHEGPPRAPSGY 1096  
Db 1063 LVSDLRIFLPPTINLDPYLRKVLKEDQQALEDGEGVYGTNALMGEPLP----- 1110  
Qy 1097 SQPPSVCSSTSFNGPFAGGVVSPQ---HSSYYSG-----MTGPQHPFYNRSGSPAP 1145  
Db 1111 -KPGSV-----GGPITTEHPHEQHTLGSNRSRSTTTGALQKSSNGLATSP 1160  
Qy 1146 GPV-----VLLNSLNVDAVCEKLEGLDQSMPLPOYCTTIKKANINGRVLAQCNIDEL 1199  
Db 1161 -PIDIDLSNVQLSSLTVEQLIELLEKVNLDKPM-ERTAPILRENAISGRVLMYCNLEEL 1218  
Qy 1200 KKENMHFGDWHLFRSTVLEMRNA-----ESHVVEDP 1232  
Db 1219 KSVLRLSFGHWEMFKLLVLTALREASVTQPASRKLSTKTTSFAGKTNDSVEMQEPQAOSTP 1278  
Qy 1233 RFLSESSSGPAPHGEPARRASHNEL 1257  
Db 1279 PFGSSTSS-----FQPIRQKSONLL 1298

Search completed: February 5, 2005, 19:34:11  
Job time : 202.5 secs

This Page Blank (uspto)